

## Durham E-Theses

---

# *ON JUDGEMENT: PSYCHOLOGICAL GENESIS, INTENTIONALITY AND GRAMMAR*

KIRKBY, DAVID,JAMES

### How to cite:

---

KIRKBY, DAVID,JAMES (2013) *ON JUDGEMENT: PSYCHOLOGICAL GENESIS, INTENTIONALITY AND GRAMMAR*, Durham theses, Durham University. Available at Durham E-Theses Online: <http://etheses.dur.ac.uk/9451/>

### Use policy

---

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full Durham E-Theses policy](#) for further details.

---

Academic Support Office, Durham University, University Office, Old Elvet, Durham DH1 3HP  
e-mail: [e-theses.admin@dur.ac.uk](mailto:e-theses.admin@dur.ac.uk) Tel: +44 0191 334 6107  
<http://etheses.dur.ac.uk>

# **ON JUDGEMENT: PSYCHOLOGICAL GENESIS, INTENTIONALITY AND GRAMMAR**

A thesis submitted for the degree of

Doctor of Philosophy

by

David James Kirkby

Department of Philosophy

University of Durham

2013

# Abstract

---

This thesis explores conceptions of judgement which have been central to various philosophical and scientific traditions. Beginning with Hume, I situate his conception of judgement within his overarching constructivist program, his science of man. Defending Hume from criticism regarding the naturalistic credentials of this program, I argue that Hume's science of man, along with the conception of judgement which is integral to it, is appropriately understood as a forerunner to contemporary cognitive science. Despite this, I contend that Hume's conception of judgement prompts a problem regarding the intentionality of judgement – a problem which he does not adequately address.

In the second part of my thesis I show how the intentionality problem which Hume grapples with is also crucial, constituting a point of departure, for Kant's transcendental undertaking. Following Kant's reasoning, I illustrate how an original concern with this intentionality issue leads Kant to a distinct conception of judgement, according to which concepts only exist in the context of a judgement. Having arrived at Kant's conception of a judgement, the remainder of the thesis is devoted to the issue of judgement forms. Kant's postulation of these forms is closely related to his conception of judgement, and I seek to establish both how these forms ought to be understood and how they might be derived. In relation to this latter issue, I suggest that there may a role for contemporary work in Generative Grammar. Specifically, I suggest that it may be viable to understand the forms of judgement as grammatical in nature, thereby securing an interdisciplinary connection between a philosophy of judgement and the empirical investigation of grammar.

## Acknowledgements

---

First thanks must go to Wolfram Hinzen. From the outset, his energy, enthusiasm and unalloyed passion for grammar have been infectious. He has been an exceptional supervisor in every respect and what's more, a good friend. Elisabeth Schellekens was instrumental in the early stages of this process in helping me clarify what it was I wanted to say. Thanks also to the Arts and Humanities Research Council who generously funded this thesis.

Doctoral life can be a lonesome business, but the company of the Language and Mind group in Durham saved me from this fate. Thanks to Tom, Jimmy, Andrew, Pallavi, Alex and Uli for tolerating my rather abrasive style, and for your feedback on my disparate collection of thoughts. Our meetings were a real highlight for me: it was fortunate indeed that we spoke the same language.

The originality of John Mikhail's work has been an inspiration for me, and I am indebted to him for his willingness to discuss it with me, as well as for the encouragement he has offered me with respect to my own. Thanks also to Nirmalangshu Mukherji for guiding my early, tentative steps towards a PhD proposal, to Nick Zangwill for taking the time to meet with me when I was struggling, and to Emma Borg and David Landy for both reading over essay drafts. I have benefitted from discussions with Beatrice Longuenesse, Paul Pietroski, John Collins, Anders Holmberg, Noel Burton-Roberts, Noam Chomsky and Amanda Taylor-Aiken.

Finally, the support of my family has sustained me throughout this process. The understanding my wife has shown has gone far beyond what I could have hoped for, while her constant willingness to lend a sympathetic ear has been all the more extraordinary given that she seldom understood what I was talking about.

# Table of Contents

---

Abstract .....	2
Acknowledgements .....	3
List of Figures and Tables .....	7
Notes on Sources and Abbreviations .....	8
Introduction .....	9
I.    A Science of Man: The Genesis of Judgement .....	15
1. Hume's Naturalistic Approach .....	16
1.1 Aims and method .....	16
1.2 Fundamentals .....	21
1.3 Against judgement-as-predication .....	25
1.4 A worry about the theory of ideas .....	29
1.5 Another worry about the theory of ideas .....	34
2. Applying Hume's Naturalistic Approach .....	45
2.1 Causation .....	45
2.2 External bodies .....	53
2.3 The thread from Hume's science of man to the <i>Critique of Pure Reason</i> ...	61
3. Cognitive Science .....	65
3.1 Picking up humean strands .....	65
3.2 Grammaticality judgements .....	73
3.3 Moral judgements .....	75
II.   Kant: The Intentionality of Judgement .....	83
4. The Nature of Experience .....	85
4.1 From synthetic a priori judgements to the possibility of experience .....	85
4.2 Relation to an object .....	91
4.3 The discursivity thesis .....	98
4.4 Global experience .....	108
5. Rules and Normativity .....	111

5.1	The normativity of the intentional .....	111
5.2	The priority principle .....	125
5.3	Objecting to Conceptualism: concepts and schemata.....	132
5.4	At the crossroads: meta-concepts.....	147
6.	The Forms of Judgement .....	152
6.1	The guiding thread of the Metaphysical Deduction.....	152
6.2	Reappraising the forms of judgement .....	156
6.3	Why Kant needs judgements of perception .....	162
6.4	There are no judgements of perception.....	164
7.	The Role of Grammar: Sketching a Research Program.....	177
7.1	Identifying the forms of judgement .....	177
7.2	Linguistic variation .....	182
7.3	Un-Cartesian linguistics.....	187
	Conclusion .....	200
	References.....	202

I confirm that no part of the material contained in this thesis has previously been submitted for any degree in this or any other university. All the material is the author's own work, except for quotations and paraphrases which have been suitably indicated.

The copyright of this thesis rests with the author. No quotation from it should be published without his prior written consent, and information derived from it should be acknowledged.

David James Kirkby



# List of Tables and Figures

---

Figure 1 .....	41
Figure 2 .....	41
Figure 3 .....	81
Figure 4 .....	100
Figure 5 .....	103
Figure 6 .....	119
Figure 7 .....	129
Figure 8 .....	154
Figure 9 .....	188
Table 1 .....	118
Table 2 .....	148

# Notes on Sources and Abbreviations

---

## Works of Hume

References to the *Treatise* will include the book, part and section number, along with the relevant page number of the cited edition e.g. ‘*Treatise*: 1.2.III, 10’.

*Treatise*                      *A Treatise of Human Nature*. 1739. (New York: Barnes & Noble Inc., 2005).

*Enquiry*                      *An Enquiry Concerning Human Understanding*. 1748. (Oxford: Oxford University Press, 1999).

## Works of Kant

In keeping with Kantian scholarship, references to the *Critique of Pure Reason* will mark the 1781 edition with the letter ‘A’, and the 1787 edition with the letter ‘B’. Paragraph number will follow letter.

*Corr*                      *Philosophical Correspondence*. 1749-1800. ed. and trans. A. Zweig. (Cambridge: Cambridge University Press, 1999).

*CPR*                      *Critique of Pure Reason*. 1781. trans. W. S. Pluhar (Cambridge: Hackett Publishing Company, 1996).

*FS*                      *The False Subtlety of the Four Syllogistic Figures Proved*. 1762. ed. and trans. D. Walford (Cambridge: Cambridge University Press, 1922).

*Logic*                      *Logic*. 1800. trans. R. S. Hartman and W. Schwarz (Indianapolis: Bibbs-Merrill Company Inc., 1974).

*Meta*                      *Metaphysical Foundations of Natural Science*. 1786. trans. E. B. Bax (London: George Bell and Sons, 1909).

*Prol*                      *Prolegomena to any Future Metaphysics that will be Able to Come Forward as Science*. 1783. trans. E. B. Bax (London: George Bell and Sons, 1909).

# Introduction

---

Suppose one wants to reflect upon the nature and character of judgement. As we might say, suppose one is interested in providing a philosophy of judgement. Where must such an inquiry begin? As Wayne Martin observes in the opening to his *Theories of Judgement*, what is required, it seems, is a conception of judgement. That is, what is required is some way of determining what one is reflecting upon and about: “to embark on an investigation of these matters we require some initial characterisation of judgement” (Martin, 2006: 1). This determination is supplied by a *concept* of judgement.

For Martin, the relevant concept of judgement is a commonly held, pre-theoretical ‘folk concept’: the range of phenomena falling within the purview of an inquiry into judgement is determined by our common sense intuitions and ordinary language. On this basis, Martin announces (2006: 1-2) that judgement is “everywhere in human life”. For example, it is involved in everything from salting one’s soup and sorting the mail to looking for a book on a library shelf and adjusting one’s course when cycling at night as obstacles appear in the headlight. As Sluga (2008) approvingly notes, for Martin, judgement is a “hyper complex” domain in the Wittgensteinian sense.

However, the problem with hyper complex domains is that, well... they are *hyper complex*! Part of what it means to call a domain hyper complex is that it resists a philosophical distillation of its nature and character. Yet this is precisely what the inquiry was supposed to be aiming at. The point here is not the completely general one that our folk concepts are open textured. Rather, it is that if we adhere to our common sense intuitions, judgement is utterly pervasive, rendering it intractable to reflective codification. Worse, not only does a philosophy of a hyper complex domain appear hopeless, it is also rather pointless. For instance, consider Martin’s (2006: 1) programmatic assertion that:

“...we can hope to understand something important about ourselves (perhaps ultimately about the kind of *being* we are) if we can understand what it is to judge”

It is hard to know what to make of this. If judgement is everywhere in human life then this assertion is not far short of tautologous. That we can learn about human *being* by learning about the character and event of human life is no doubt true, but it is hardly informative.

More to the point, if judgement is everywhere in human life, we would do just as well to investigate human life outright and be done with it. The concept of judgement does not seem to add anything.

As such, if we adhere to our folk intuitions and ordinary language, the domain of judgement seems far too nebulous, expansive and unwieldy to be the object of philosophical inquiry. Perhaps this is why the concept of judgement is, philosophically, a little antiquated (Glock, 2010: 1). Perhaps a thesis on judgement is not a good idea.

On the other hand, conceptions of judgement *have* played a prominent role in various philosophical and scientific traditions. For example cognitive scientists commonly speak of investigating subjects' *judgements*. Furthermore, the claim that experience is fundamentally *judgemental* is one of the central claims made by Kant in the *Critique of Pure Reason*. Notice however, the plural marking: conceptions of judgement. There are two tightly related points here. Firstly, where conceptions of judgement have been prominent or theoretically significant, their contours have not been determined by recourse to a common pool of ordinary language and folk intuitions. These conceptions of judgement are not interchangeable with, and have not been intended to be interchangeable with, a folk concept. As such, although our folk or pre-theoretic conception of judgement may indeed denote a hyper complex domain, generally speaking, philosophical conceptions of judgement have not converged upon this domain. Instead, and this brings me to the second point, where conceptions of judgement have been prominent, judgement has been understood as performing a function or role *relevant to the theoretical undertaking in question*. Conceptions of judgement have been prominent where they have captured a theoretically salient function of this kind. As such, far from being determined by common sense intuitions, judgement has been understood in technical or theory-laden terms - motivated by, and answerable to, the cluster of problems and questions constitutive of the inquiry at hand. As a consequence however, different philosophical undertakings and approaches have yielded different conceptions of judgement. Not only do conceptions of judgement not converge upon a hyper complex domain, they do not converge at all.<sup>1</sup>

---

<sup>1</sup> Of course, this does not entail that conceptions of judgement have been completely unrelated to our pre-theoretic conception of judgement. Were this the case, it would beg the question as to why the term 'judgement' is being used at all. The important point however is that as technical concepts, they are not answerable to or determined by our common sense intuitions and ordinary language. For instance, consider the relation between the technical concept of language employed by generative linguists and the folk concept of language. Although

For this reason, speaking of *theories or philosophies* of judgement is misleading, since it implies that ‘judgement’ is un-problematically univocal, denoting a fairly uniform range of phenomena which is the subject matter of various and competing theories. This implication is precisely what I am here denying. In this thesis therefore, it is not my aim to provide a theory of judgement, but rather, to situate conceptions of judgement within the theoretical contexts within which they make sense and against which they can be evaluated. Specifically, I shall concern myself with two distinct conceptions of judgement, central to Hume’s science of man and Kant’s transcendental philosophy respectively. The challenge will be to show how these fit into the relevant philosophical and scientific undertakings, and the role they play therein.

Unfortunately, while this approach succeeds in eschewing an inherently nebulous and unwieldy object of inquiry in favour of much more tractable subject matter, it comes at a significant cost. Specifically, insofar as we are dealing with two disparate conceptions of judgement, in what sense is a discussion which moves from one to the other philosophically continuous? What ensures that such a discussion is more than fragmented, juxtaposed, descriptions of how judgement has been variously understood? Something must serve to unite or bind together the relevant conceptions of judgement if a discussion of them is to be philosophically continuous. Yet, short of casting ourselves back into a domain of hyper complexity, it cannot be judgement which imparts this unity. Perhaps a thesis on judgement is really not a good idea.

This issue strikes at the heart of this thesis. From its most nascent beginnings, the overarching aim of the thesis has been to explore new possibilities for bringing philosophy and the empirical sciences of the mind to bear upon each other, new ways in which each could inform the progress of the other. Since conceptions of judgement feature both in cognitive science and in various philosophical traditions as the bearer of epistemological and alethic properties, it originally appeared to be a promising commonality in this respect: a prism through which connections could be drawn between the two disciplines. However, as I realised that the conceptions of judgement being utilised were theory-laden, these prospects dimmed as a result. For example, if the conception of judgement utilised by cognitive scientists suffices for and serves the explanatory goals definitive of the field, what is the relevance of, for example, Kant’s pronouncements regarding judgement, bound up as they are

---

as a technical concept, it is unconstrained by common sense intuitions about what ‘language is’, the use of the term implies a certain amount of overlap.

with his transcendental method? Well, *prima facie*, arguably none. Insofar as Kant's conception of judgement is shaped by very different issues and problems from those concerning cognitive scientists, it is hard to see how this could constitute a productive engagement of philosophy and cognitive science. Of course a juxtaposition of these conceptions of judgement could still be interesting, and each could still be assessed according to their own standards of evaluation; yet juxtaposition falls short of comparison, and in the absence of comparison, it is hard to see how such a discussion could illuminate and/or strengthen the relationship between cognitive science and philosophy.

So far, so discouraging, and were this the end of the matter, I would have had to either discard the topic of judgement altogether or curtail my philosophical ambitions. Fortunately however, the matter does not end there. What I noticed is that the conception of judgement central to cognitive science is intimately bound up with a challenge regarding the possibility of being in an *intentional state*. This challenge is more accurately described as a problem, I call it the 'intentionality problem', since it follows directly from cognitive scientists' own conception of judgement, and is thorny indeed. From the perspective of this thesis, the reason this is significant is because this intentionality problem is also central to *another* conception of judgement: Kant's. In fact, Kant's conception of judgement is only comprehensible in relation to this problem: they are inextricable entwined. It follows from this that a discussion of these two conceptions of judgement *will* be philosophically continuous after all. Insofar as one conception inevitably runs up against this problem, and the other is inextricably bound up with it, intentionality serves to impart unity upon the discussion.

The chapters of this thesis are divided into two parts. Part I, encompassing chapters one to three, concerns a conception of judgement which I trace back to Hume, identifying it as integral to a naturalistic tradition which, following Hume, I label a 'science of man'.

Chapter One outlines the philosophical foundations of Hume's science of man, identifying the conception of judgement integral to this approach.

Chapter Two describes some philosophical results of Hume's naturalistic approach. I show how the relevant conception of judgement inevitably leads to the intentionality problem (or as Hume puts it, the problem of how we come to believe in external objects). I argue that one of Hume's greatest failures in the *Treatise* is his attempt to grapple with this problem, proffering as he does a completely circular account.

Chapter Three explains how Hume's science of man is nowadays pursued by cognitive scientists. Importantly, this continuity is marked by a conception of judgement common to that of Hume.

Part II, encompassing chapters four to seven, broadly concerns Kant's conception of judgement. Specifically, it is claimed that Kant's conception of judgement only makes sense in light of his concern with the intentionality problem (or, in his terms, the possibility of objective cognition or relation to an object).

Chapter Four locates the intentionality problem within the context of Kant's overarching transcendental philosophy.

Chapter Five discusses Kant's response to, or account of, the intentionality problem. According to Kant, intentionality presupposes rules, and rules presuppose judgement. As such, Kant's conception of judgement is inextricably bound up with the role it plays in making intentionality possible.

One way of (rather drastically) simplifying the dialectic up to this point would be as follows: I begin with the science of man conception of judgement and then show how this conception gives rise to a problem regarding intentionality; subsequently, I turn to Kant's recognition of this intentionality problem and then show how Kant's conception of judgement is part of the solution to this problem. In other words, while two conceptions of judgements are discussed, they are not directly related. Rather, they are related mediately, by both being bound up with the problem of intentionality.

In the final two chapters of the thesis, chapters six and seven, I seek to develop, or rather, amend a central element of Kant's conception of judgement: the nature and role of the *forms* of judgement. While in one respect these two chapters are absolutely continuous with what goes before – the conception of judgement which is being amended is largely outlined in the preceding chapters – they do stand somewhat apart from the primary, sweeping dialectic of the thesis, concerned as it is with connecting two disparate conceptions of judgement via the intentionality problem. By contrast, chapter six and seven have a more narrow and focused topic: how one aspect of Kant's conception of judgement, specifically, the role of judgement forms, ought to be understood. I argue that although Kant was right to think that there are such forms, and although he was right to think that they are of

philosophical significance, he mistakes their philosophical significance and has no sound means for identifying them.

Chapter Six concerns the philosophical significance of the forms of judgement. I argue, *against* Kant, that any such forms are appropriately understood as meta-concepts, and as such, as foundational for objective cognition.

Chapter Seven addresses the issue of how the forms of judgement are to be identified. I argue that these forms ought to be derived from our best theories of *grammar*. That is, grammar (a universal and uniform property of the human species) provides the table of judgement forms, understood in Kant's sense. Although this claim is not decisively established, requiring as it would a thesis in its own right, I nevertheless seek to show that it is both plausible and attractive: amounting to a viable research program.



## Part I

# A SCIENCE OF MAN: THE GENESIS OF JUDGEMENT

If I were to judge that Sellars was the greatest philosopher of the twentieth century in a room full of philosophers, controversy is likely to ensue. Possibly, some of my interlocutors would agree; probably, many would disagree, citing the brilliance of Wittgenstein, the importance of Frege, or, indeed, the opacity of Sellars' own writing. Irrespective of the precise arguments, the controversy would turn upon the question of whether I judged well or not.

Suppose that this debate were witnessed by a passive observer, passive in the sense that she is utterly disinterested in Sellars' philosophical credentials. Instead of following the debate, what she is observing is *me*: what I do, what I say etc; she observes my behaviour because she wants to understand why I behave in the ways that I do.<sup>2</sup> As such, what interests her is not the correctness of the judgement, but the judgement itself or rather, *the fact that I made it*. In her eyes, my judgement is part of the data set which her theory needs to accommodate, in the same way that it must accommodate other behavioural qualities which I exhibit.

This part of my thesis is concerned with a tradition, beginning with David Hume and upheld in the modern era by cognitive scientists, in which judgement is understood and treated as by the passive observer. Having observed the judgements people make in the same way that one might observe other behavioural qualities, the challenge, according to this tradition, is to explain these judgements by tracing their *genesis*. Following Hume's lead, I shall call this tradition the science of man.

---

<sup>2</sup> That such a theory is plausibly beyond the reaches of human understanding – it would amount to what Noam Chomsky calls a 'theory of everything' – is irrelevant to the point.

# Chapter 1. Hume's Naturalistic Approach

---

My aim in this chapter is not to *vindicate* a naturalistic reading of Hume – alternate interpretations will be addressed only indirectly – but to work ‘within’ this reading of Hume. In particular, I shall underline the internal consistency and coherence of Hume’s naturalism by (a) distinguishing it from different philosophical enterprises, and (b) defending it from some common objections.

## 1.1. Aims and method

Hume begins *A Treatise of Human Nature* by stating his intention to pursue a science which will “introduce the experimental method of reasoning into moral subjects”.<sup>3</sup> ‘Moral subjects’ here denotes simply human beings, rather than a morally praiseworthy subset thereof (Flew, 1986: 2). In explicitly targeting human beings, Hume means to demarcate the explananda of his inquiry from those of other natural sciences which have a much broader purview. Biology or physics, for example, provide descriptive laws and generalisations which are applicable across the natural world and, for that reason, have nothing specifically to do with moral subjects. Of course, physics stipulates the rate at which a man will accelerate having stepped off the edge of a cliff, but it does so in the same way for any object, animate or inanimate. In contrast to this, Hume is interested in “those respects in which they [human beings] differ from other ‘objects of nature’” (Stroud, 1977: 2). The inquiry is directed at those qualities which are unique or specific to human beings, or at least those which appear to be so. Rate of acceleration is not such a quality. Instead, Hume chooses to focus on certain beliefs in matters of fact which people form. The beliefs he targets are fundamental to our day to day activity and reasoning, and are apparently human universals. Although Hume tends to speak of belief, he finds no distinction between this and judgement and uses the terms interchangeably (*Treatise*: 1.3.VII, fn. 1; cf. Owen, 2003: 17 & fn. 2). My discussion of his views shall follow him in this respect.

While Hume’s inquiry has a much narrower subject matter from that of other natural sciences, it adopts the same experimental method of reasoning. Clearly, this method was

---

<sup>3</sup> My discussion of Hume will be mostly based upon and drawn from the *Treatise*, particularly Book 1.

shorn of the benefits that technological advancement has brought to experimentation in the current age. Nevertheless, at the heart of all experimentation is the careful observation of how phenomena co-vary in systematic ways with changes in relevant circumstantial and situational factors. Such careful observation, Hume hoped, would reveal the formation mechanisms responsible for our judgements. The task was to trace the genesis of as many judgements as possible back to as few causes as possible, thereby delivering a principled or *explanatory* account of the explananda.

This model of explanation or intelligibility is perfectly familiar. Ensuring that “all our principles are as universal as possible, by tracing up our experiments to the utmost, and explaining all effects from the fewest and simplest causes” (*Treatise*: intro., xxvi) is integral to any scientific inquiry, and it is no coincidence that Hume called his investigation simply a ‘science of man’. Ever since Smith’s famous papers (1905 & 1905a), readings of Hume that have placed this science of man at the heart of his philosophy have been labelled ‘naturalistic’.<sup>4</sup> According to this naturalistic reading, Hume is first and foremost interested in instigating a positive, scientific program of inquiry rather than propounding sceptical conclusions or advancing empiricist theses.<sup>5</sup> In other words, it takes seriously Hume’s opening statements at the beginning of the *Treatise* about wanting to apply the experimental method to moral subjects, taking this to be the touchstone of his whole philosophy; his naturalistic approach *is* his science of man. Key to this science is a distinction between explanation and justification:

“Hume’s recommendation is to replace endless and fruitless “cogitating” in an attempt to give a philosophical justification of our beliefs, with an attempt to find a scientific explanation of their origin.” (Biro, 1993: 44-5)

Just as the passive observer seeks to account for why I make the judgement ‘Sellars is the greatest philosopher of the twentieth century’ without concern for whether the judgement is correct, so too Hume wants to know *why* we form the judgements that we form, not whether they are correct or not. As a consequence of this, a science of man proceeds in abstraction from any normative assessment of our beliefs. By dispensing with the question of the

---

<sup>4</sup> There is (at least) one other respect in which Hume is commonly said to be a naturalist. Hume argues that many of our beliefs are *natural* in the sense of being inevitable and indispensable (Smith, 1905: 152; Norton, 1982: 15-19). Nature has not left us a choice in these matters. Although this is absolutely compatible with the naturalism of a science of man, neither type of naturalism entails the other, and so they should be held apart.

<sup>5</sup> Of course, this reading of Hume is perfectly consistent with the presence of empiricist or sceptical arguments in the *Treatise*; these elements are simply viewed as secondary to his naturalistic aims.

correctness of our judgements, so too Hume dispenses with the evidential or justificatory standards according to which such correctness is established. The justificatory basis of our judgements is a matter left to others. For example, when Hume famously asks of inductive inferences “whether we are determined by reason to make the transition, or by a certain association and relation of perceptions” (*Treatise*: 1.3.VI, 73), he is posing a question of cognitive psychology, not a normative one about their epistemic warrant (Allison, 2008: 112).<sup>6</sup>

It is salutary in this respect that Hume includes in the *Treatise* a chapter on the reason of non-human animals. Indeed, he declares that “when any hypothesis, therefore, is advanced to explain a mental operation, which is common to men and beasts, we must apply the same hypothesis to both” (*Treatise*: 1.3.XVI, 138). This is actually a very strong claim and it is far from clear that Hume is entitled to it. Setting this aside however, the quote serves to illustrate the parity of aims and explanatory standards Hume sees between his science of man on the one hand, and sciences of non-human animals on the other. Consider a couple of well catalogued facts about frogs. Frogs will attack any object the size of an insect or a worm, its natural prey, provided that the object is made to move in a sufficiently convincing manner. Similarly, frogs reflexively extend their tongue towards both flies, and objects resembling flies. These are distinctive characteristics and ones which any science of the natural dispositions and behaviour of a frog would need to target. Scientists are interested in explaining this behaviour and this is achieved, at least in part, by investigating the kind visual image generated from light striking the retina. Essentially, the challenge is to explain “what the frog’s eye tells the frog’s brain” (Lettvin, et al., 1959). Crucially however, this explanation is distinct from an assessment of when, if at all, a frog’s behaviour is *correct* or *successful*. Of course, criteria for such an assessment quickly suggest themselves. For example, a frog’s act is a success if it catches an insect (still though, will any insect do? what if frogs are unable to digest some insect species?) and unsuccessful if it catches an object which merely resembles an insect. The point, however, is that such assessment need not be addressed by an investigator targeting the causal genesis of the behaviour: a scientist of a frog, or of any other animal, can construct adequate explanatory accounts of the relevant behaviour without deigning to consider whether the behaviour is successful. So too in the case of

---

<sup>6</sup> Of course, there is a relationship between the kind of explanatory cognitive psychology Hume is interested in and the justificatory considerations an epistemologist might cite. After all, attributing a belief to certain formation mechanisms may help determine whether it is justified or not. However, the point is that this question of justification is something over and above the explanation itself and one can investigate the latter whilst remaining neutral with regards to the former.

humans, establishing the origins of our judgements can be addressed without determining whether they are correct:

“...belief-producing processes were [for Hume] neutral with respect to whether the beliefs produced are true or false” (Hatfield, 1990: 26; cf. Norton, 1982: 201-208).

Hume is not at all reticent when it comes to stating the inestimable value of this naturalistic approach. It is extraordinary, he says, that this matter has not received more attention given that all other sciences rely, to some extent, upon the powers and faculties which underlie mans’ capacity for understanding, reasoning and judging. Although some of the “late philosophers in England” may have had a hand in its development (*Treatise*: intro., xx-xxi), Hume clearly feels that it has not received anything like the attention that it deserves, claiming that the foundation of this science will be “almost entirely new” (*Treatise*: intro., xvi).

The neglect of this science of man was largely a product of the traditional way of conceiving human beings. Ever since the ancient Greeks, man had been understood as a fundamentally rational creature. In both the practical and theoretical spheres, human activity was thought to be governed, primarily, by the apprehension of relevant reasons.<sup>7</sup> Judgements, the ‘behavioural qualities’ which would interest Hume so much, were delivered by the intellect and relied upon rational insight. Whether or not this was deemed to mark the metaphysical essence of man or not, it had significant implications for what a philosophy of man could be. In contrast to beasts who, acting only according to the dispositions of their organs, could be studied as complex machines (Descartes, 1952: 20), man would not admit of such an inquiry. Human beliefs and actions were a matter of the reasons available to the agent at the time and so “purely intellectual” (Smith 1905: 165). As Stroud nicely puts it, “any ‘explanation’ of the beliefs or actions of a distinctively rational agent could therefore do no more than show that, on the evidence then available to him, those beliefs or actions were the most reasonable ones for that agent to adopt” (Stroud, 1977: 13). Such an ‘explanation’ however, would rest upon a philosophical account of what a good reason is, not upon the kind of descriptive generalisations delivered by an experimental approach. This foreclosed the possibility of a science of man.

---

<sup>7</sup> This is an oversimplification. It ignores important issues like the possibility of a weak will: can a man fail to will what he recognises he has (sufficient or overriding) reason to do?

Indeed, this partly accounts for the pre-eminence of epistemology in the work of Hume's predecessors, both empiricist and rationalist alike. Philosophical work was required not to explain why men behave in the ways that they do – an unfeasible task - but to establish secure epistemic ground for such behaviour. That is, how can reasons be secured as *good* reasons? For example, Descartes begins *Rules for the Direction of the Mind* with the programmatic assertion that “the end of study should be to direct the mind towards the enunciation of sound and correct judgements on all matters that come before it.” The main prong of this enterprise consisted of showing that thought is appropriately related to the external world and given this it was all but inevitable that scepticism would be a threat which demanded attention and rebuttal.

Now, it is *not* the case that adopting the experimental method with respect to moral subjects presupposes, either conceptually or methodologically, that the vision of man as a rational being is false.<sup>8</sup> It is perfectly possible to establish empirically that people adopt the beliefs that they do according to reflection upon relevant reasons.<sup>9</sup> Although this finding would not show that these reasons are valid or correct – a ball that is only ever in the epistemologists' court – it would reaffirm the rationalistic picture of man governed by the apprehension of reasons. What marks out a science of man is not that it assumes the falsity of this picture, but that it refuses to accept its accuracy *a priori*. Everything should be subjected to careful observation and experimental method, including how our judgements are delivered. Our judgements may be produced by the apprehension of reasons, but on the other hand they may be produced via “secret springs and principles” of the mind (*Enquiry*: 1.15, 93), analogous to the ‘occult’ forces Newton had reintroduced to nature. Moreover, since it is possible that reason be efficacious with respect to some judgements, but not others, this method will have to be applied on a case by case basis.<sup>10</sup> As such, Hume's naturalistic approach involved a disavowal of any claim about the essence of the mind (Hatfield, 1990: 26).

---

<sup>8</sup> I think that this point is often lost amidst the numerous remarks Hume famously makes regarding the subordinate role of reason: that it is a slave to the passions etc. However, such remarks ought to be understood as a conclusions or even findings of Hume's inquiry, rather than bound up in the nature of the inquiry itself. Adopting a science of man does not entail that the rationalist vision of man is false, though it may lead to such a conclusion.

<sup>9</sup> Kohlberg's psychological account of moral judgement is an example of an empirical approach supporting rationalism in this way (Kohlberg, 1981)

<sup>10</sup> Indeed, although Hume is renowned for concluding that, in matters of belief, reason is subordinate to our natural propensities, even he accepts that reason produces *some* of our beliefs: arithmetic relies upon demonstration i.e. proofs. In the *Enquiry* he includes also geometry under this heading. Whether reason is also responsible for those relations of ideas which are intuitive as opposed to demonstrative, e.g. ‘white is not black’, I set aside (Smith thinks that they are, calling them the products of “rational necessity” (1905: 157)).

## 1.2. Fundamentals

The theory of ideas states that to be in a particular mental state is to have a mental particular or an array of such particulars ‘in mind’. Any type of experience which is continuous over time is constituted by a succession of these reified mental objects. Allison calls this model of thought that of ‘seeing with the mind’s eye’ (Allison, 2008: 6). This model was adopted by Hume’s empiricist predecessors such as Locke and Berkley, and it is clear that he followed them in this respect.

“We may observe... that nothing is ever really present with the mind but its perceptions or impressions and ideas and that external objects become known to us only by those perceptions they occasion. To hate, to love, to think, to feel, to see; all this is nothing but to perceive.” (*Treatise*: 1.2.VI, 56)

In the context of Hume’s science of man, his theory of ideas is a “fundamental assumption” (Smith, 1905: 159, fn. 2). This is because it constitutes the theoretical framework within which he seeks to explain our judgements. ‘Establishing their genesis’ means tracing them back to their origins within this mental domain.

Hume calls all constituents of the mental domain “perceptions”. All perceptions, with the exception of the passions, have content; that is, they have accuracy or veridicality conditions (Garrett, 2006: 302-306). The most prominent classificatory divide amongst them is that of impressions and ideas. Impressions form a very broad category, including sensations, emotions and passions. The essential mark of an impression, that which properly distinguishes it *as* an impression, is the force or liveliness or violence (Hume uses the terms interchangeably and, following Allison, I shall abbreviate them to ‘FLV’) with which they impose themselves upon consciousness.<sup>11</sup> Impressions constitute that part of our mental life called ‘sensing’ or ‘perceiving’. For example, that I find myself confronted by a laptop, with a fragment of a PhD thesis on the screen is determined by the array of impressions I have ‘in mind’. In other words, impressions constitute that sensory stream to which a subject is only a

---

<sup>11</sup> Martin (2006: 27-28) notes that Hume’s use of FLV is ambiguous. It could denote either a phenomenological quality of how percepts ‘appear to the mind’s eye’ i.e. vividly or clearly etc, or it could denote the psychological ‘momentum’ of the percepts, which better fits the use of the term ‘force’. David Landy (2008) urges that FLV should be understood in the latter sense: functionally, as concerning the *behaviour* of perceptions. My view is that the most natural way to read Hume is that while percepts with greater FLV do function in a certain way, this is *because* they exhibit especially vivid phenomenological qualities.

passive recipient. In Kantian terms, the capacity for impressions is purely receptive. Short of closing my eyes or turning around, I cannot dispel the impression of my laptop and, in turn, if I do close my eyes, I cannot conjure the impression through my own volition. What I would conjure in this case would not an impression, but the *idea* of a laptop. Ideas, for Hume, are *images* which are utilised in thought, reasoning and language (*Treatise*, 1.3.X).<sup>12</sup> As the essential mark of impressions is the amount of FLV they exhibit, so, correlatively, ideas are distinguished by their relative lack of FLV. They are images which are ‘fainter’ or ‘less vivid’ than impressions.

There are three important points to make about Hume’s distinction of impressions and ideas. Firstly, although they are distinguished, they are not so on the basis of type or kind; the difference resides only in the ‘amount’ of FLV they exhibit. Indeed, Hume notes that although the line between impression and idea is generally a clear one, “in particular instances they may very nearly approach to each other” (*Treatise*: 1.1.I, 7). Secondly, since it is impressions and ideas which characterise sensing and thinking respectively, this means that the difference between sensation and thought is similarly one of *degree*, not kind. Thirdly, while Hume’s characterisation of perceptions in terms of FLV was not an innovation (e.g. Berkeley, 1710: 78), what did set Hume’s view apart was how the FLV of a perception was taken to be *no guide* as to its correctness. At no point do we find Hume claiming that the more lively/vivid/clear a percept is the more assured we can be of its veridicality. The temptation to draw epistemic conclusions (or even indications) from the clarity of a percept was one which Hume assiduously resisted – as would be expected of a philosopher primarily interested in a science of man. By contrast:

“I came to the conclusion that I might assume, as a general rule, that the things which we conceive very clearly and distinctly are all true.” (Descartes, 1637: III, 52)

Alongside the impression/idea dichotomy, Hume employs another which cuts across it. Both impressions and ideas can be either simple or complex. Simple impressions and ideas admit of no distinction or separation because they have no constituents. For example, one might have an impression of an apple, but it must be complex for it can be separated into components pertaining to its taste, colour and smell etc. Similarly for ideas, they are complex if they can be broken up into constituents, and simple if not. In other words, a percept is

---

<sup>12</sup> Arguably, Hume also allowed for a small class of non-imagistic ideas – ‘relative ideas’ (see Flage, 1990: 42–51).



complex if and only if it is decomposable, otherwise it is simple. This atomistic account of mental particulars is governed by what is called in the literature Hume's 'copy principle':

"...all our simple ideas in their first appearance are derived from simple impressions, which are correspondent to them, and which they exactly represent" (*Treatise*: 1.1.I, 9)

This principle involves two theses:

- (i) Every simple idea resembles an impression

"The first circumstance, that strikes my eye, is the great resemblance betwixt our impressions and ideas in every other particular, except their degree of force and vivacity." (*Treatise*: 1.1.I, 8)

For every simple idea there is a resembling impression. For example, my idea 'red' resembles an impression 'red' which I have experienced.

- (ii) Every simple idea is causally dependent upon a simple impression

That some simple impressions and some simple ideas resemble each other naturally suggests a dependency relation holding between them. Hume argues that for any resembling impression-idea pair, the impression always appears in experience prior to the acquisition of the idea i.e. the impression 'red' precedes the resembling idea. This temporal ordering establishes that it must be the impression which causes the resembling idea and not vice versa.<sup>13</sup>

It is critical to underline that the copy principle is restricted to *simple* ideas and impressions. Hume realises that people can have complex ideas for which there has not been any resembling complex impression. For example, people can entertain complex ideas such as spiteful giraffes and golden mountains without having had any resembling impression. The distinction between those complex ideas which are copied from complex impressions and those which are not maps onto the distinction of the two faculties which are responsible for our ideas: the memory and the imagination. Hume takes it as evident that ideas of the memory have greater FLV than those of the imagination, but the most significant difference resides in this copying issue. Complex ideas of the memory are copied from corresponding

---

<sup>13</sup> This view places Hume firmly in the empiricist camp, since it is equivalent to the claim that there are no innate ideas (cf. Allison, 2008: 20-21).

complex impressions. Although we can recognise the simple parts of the idea, the memory does not have the liberty to separate or reorder them in ways contrary to the impression from which it was copied anymore than a historian can rearrange past events. The ideas of the memory are cast in the stone of the original impression. In this respect, the imagination could not be more different:

“...all simple ideas may be separated by the imagination, and may be united again in whatever form it pleases” (*Treatise*: 1.1.IV, 13).

What for the memory are unbreakable bonds, the imagination is able to dissolve ‘as it pleases’, decomposing any complex idea into its simple parts. From this, it is at perfect liberty to compound and recompose these materials into complex ideas of any form or order. Notice that this liberty does not amount to *creating* new simple ideas, an impossibility given that all simple ideas are copies of simple impressions, but has to do only with the way they are put together. It is this power that enables us to entertain the ideas of winged horses and monstrous giants, of which we have had no corresponding impressions. A useful metaphor for this account of the imagination is a child playing with building blocks. All complex structures are composed of simple blocks which the child can transfer from one structure to another, detach from a given structure or break up entirely and reconstruct in an original way. The only limitation is the blocks the child has to play with.

This is not a complete account of the imagination. The liberty of separating any simple ideas and reuniting them in any form, does not explain the *regularity* of our ideas. Were our ideas entirely loose and unconnected, it would be mere chance that would move us from one to the other. Hume finds it obvious that this cannot be correct:

“...even in our wildest and most wandering reveries, nay, in our very dreams, we shall find, if we reflect, that the imagination ran not altogether at adventures, but that there was still a connection upheld among the different ideas, which succeeded each other” (*Enquiry*: 3.1, 101)

There must be principle(s) of the imagination which ensure the cohesiveness or regularity of our ideas. Although they cannot be so strong as to make two ideas inseparable, for this would violate the liberty of the imagination, there must be a “gentle force” or “a kind of attraction” which pushes certain ideas together. The force is that of association. The qualities of ideas which trigger their association in the imagination are: resemblance, contiguity (spatial and

temporal) and cause and effect, which Hume characterises also as ‘natural relations’. These qualities prompt the associative connection of ideas in the imagination, serving to tie together what would otherwise be disparate objects of the mind.

### ***1.3. Against judgement-as-predication***

If the theory of ideas provides the theoretical framework for Hume’s naturalistic program, a crucial supplement to this framework is Hume’s conception of judgement. Nor should this be surprising given that an investigation of the genesis of a judgement can only be regarded as successful insofar as there is some prior apprehension of what the investigation is tracing the genesis of.

Hume’s predecessors had understood judgement to be the separation or combination of multiple ideas. For example, Locke (1690: IV.XIV, 334) considered judgement to consist of “putting ideas together or separating them from one another in the mind”, which is a matter of how “the mind takes its ideas to agree or disagree” (*Ibid*). For instance, the judgement ‘white is not black’ is a matter of recognising or “seeing” that the ideas ‘white’ and ‘black’ do not agree, in a certain respect, and so separating them.<sup>14</sup> Locke proposed four relations of agreement/disagreement betwixt ideas: identity, necessary entailment, coexistence, and real existence (Locke, 1690: IV.I, 255).<sup>15</sup> To recognise that two ideas agree or disagree in one of these respects i.e. that they are related in a certain way, *is* to unite or separate them in a judgement. Similarly, the Port-Royal logicians considered that after forming ideas, “we unite those which belong together by affirming one idea of another; we separate those which do not belong together by denying one idea of another. *To judge is to affirm or to deny.*” (Arnauld, 1662: 108) [my emphasis]. This uniting or separating of ideas corresponds to the contemporary notion of predication. For example, in the judgement,

---

<sup>14</sup> Locke’s identification of ‘uniting or separating ideas’ and ‘recognising/perceiving ideas to agree or disagree’ is confusing since the former seems to imply an activity on the part of the agent, apparently lacking in the latter case. Nevertheless, they are interchangeable for Locke. To unite two ideas *is* to recognise a relation of agreement holding of them and to recognise such a relation is to unite them.

<sup>15</sup> In Locke’s terminology, strictly speaking, only the relation of ideas by means of coexistence and real existence would be a judgement. This is because Locke draws an epistemic distinction between *judgement* and *knowledge*. As Owen (2003: 17) puts it, “judgement is the admitting or receiving of any proposition as true, without certain knowledge that it is so.” Essentially, judgement concerns those relations between ideas that can only be probable (coexistence, real existence), and so lack the epistemic security of demonstrative or intuitive knowledge (identity, entailment). However, in the context of Hume’s discussion, what is of interest is the difference between conception and judgement and consequently, it should be clear that by ‘judgement’, I denote also Lockean knowledge.

‘ducks can fly’ the idea, ‘can fly’ is predicated of ‘ducks’. According to Hume’s predecessors, judgement is a matter of predicating one idea of another (cf. Anderson, 2008: 92-94). Call this the predication theory of judgement. Hume, however, vehemently opposed this theory of judgement.

“We may here occasion to observe a very remarkable error which... is universally received by all logicians. This error consists of the vulgar division of the acts of the understanding into *conception, judgement and reason*, and in the definitions we give of them. Conception is defined to be the simple survey of one or more ideas: judgement to be the separating or uniting of one or more ideas: Reasoning to be the separating or uniting of different ideas by the interposition of others.” (*Treatise*: 1.3.VII. fn. 1)

Hume provides two arguments against the predication theory of judgement. The first concerns the matter of existential judgements. Hume argues that we do not have a clear notion of ‘existence’ as a distinct idea. In an insightful passage in which he largely pre-empts Kant’s more renowned discussion of the matter, Hume argues that to conceive of an object as existing does not add any extra quality to it (*Treatise*: 1.3.VII). It follows that in case of existential judgements, e.g. ‘God exists’, ‘God is’, ‘It is raining’, there is only one idea involved. Given that we can judge of single ideas, judgement cannot be a matter of separating or uniting ideas.

Secondly, Hume argues that even where a judgement clearly does involve multiple ideas it cannot be their predication which distinguishes it *as* a judgement. The reason for this is that the predication theory cannot adequately explain the difference between incredulity and judgement. Incredulity, for Hume, includes both the denial of a proposition, and the mere conception or supposition of a proposition and I shall treat these topics separately. Consider a case where John and Bill disagree over whether Caesar died in his bed; John judges that he did and Bill denies it. Hume points out that despite his denial, Bill clearly understands the relevant proposition and forms the same ideas as John: that Caesar died in his bed; were this not the case, it would not be a case of genuine disagreement, but of two interlocutors talking past each other. There must be some content, ‘Caesar died in his bed’, which both John and Bill converge upon, and which can be fixed independently of John’s believing it and Bill’s denying it. Wayne Martin (2006: 35) calls this ‘Hume’s identity condition’. Consequently, the possibility of disagreement shows what the difference between incredulity and judgement

*is not*: it cannot be a matter of the content grasped. As such, a distinction can be drawn between:

- (a) The content of a possible judgement/denial
- (b) A judgement<sup>16</sup>

Although judgement or denial obviously presupposes a grasp of whatever content is being judged or denied, the critical point is that it is no feature of the content which determines whether it is judged of or denied; these are separate issues: (a) is independent of (b).

However, this distinction is impossible on the predication theory of judgement. According to this, by judging that Caesar died in his bed, John is predicating ‘died in his bed’ of ‘Caesar’. On the other hand, by denying the proposition, Bill is separating the ideas ‘died in his bed’ and ‘Caesar’. However, this means that there is no common content shared by both John and Bill, no combination of ideas apprehended by both. In other words, the difference between judgement and incredulity is explained in terms of a difference in content: (a) and (b) are run together. Yet, this is precisely what the possibility of disagreement shows cannot be correct.

A parallel argument, one which does not involve disagreement, can be developed with respect to the difference between the supposition of certain content and judgement. Imagine that I begin a lecture by announcing the thesis to be defended: ‘Sellars is the greatest philosopher of the twentieth century’. It is surely possible for someone in the audience to begin by merely entertaining or supposing this thesis and then come to believe it over the course of the lecture (presumably due to the eloquence of my arguments). Yet, according to the predication theory of judgement, this is impossible. Since judgement is simply the combination of ideas, either the audience member predicates ‘is the greatest philosopher of the twentieth century’ of ‘Sellars’ at the beginning of the lecture – in which case she judges it rather than merely entertains it – or she does not, in which case there is nothing for her to entertain. As in the case of disagreement, the problem is evidently that the predication theory obliges us to view the difference between judgement and supposition as one of the content formed or grasped. Once these two issues are confounded, there are no means by which to fix

---

<sup>16</sup> There is clearly a sense then, in which Hume propounded an early version of Frege’s famous distinction between (1) the apprehension of a thought and, (2) the recognition of the truth of a thought – judgement (Frege, 1956: 294). Nevertheless, Rodl (2009: 438-439) is right to argue that this should not be overstated. For Hume, the distinction applies only to judgements of matters of fact, not those concerning relations of ideas (essentially, he preserves a predication account of judgements concerning relations of ideas: a nuance I have set aside here).

a notion of content which can at  $t_1$  be merely supposed and at  $t_2$ , judged. Indeed, in Locke's account, the notion of judgement is explicitly built into the mechanism of content formation. As already noted, for Locke, the combination or separation of ideas (predication) is a matter of *perceiving* or *recognising* that they agree with each other. Yet, in this context, 'perceiving a relation of agreement' just seems to mean 'assent or judge that ideas agree'. So judgement is entwined with predication from the outset:

“... predication isn't [for Locke] distinct from affirmation or denial; understanding propositional content isn't distinct from knowing or judging it to be true” (Owen, 2003: 17)

What Hume realised is that the difference between incredulity and judgement cannot be one of content. This essentially presented him with a wholly original question, one necessarily neglected by his judgement-as-predication predecessors: “wherein consists the difference betwixt incredulity [i.e. predication] and belief?” (Treatise: 1.3.VII, 77-78).

Hume's answer to this question is disarmingly simple. Since the difference between judgement and incredulity cannot reside in the content of ideas, Hume concludes that it must only be the *manner* in which ideas are conceived which separates belief from incredulity. A belief is simply a strong and lively idea. To believe that the sun is shining is to 'vividly' picture the sun shining.<sup>17</sup> In this way, judgement is not a distinctive product of the understanding, but is properly resolved into conception (Treatise: 1.3.VII, fn. 1).

“When I am convinced of any principle, 'tis only an idea, which strikes more strongly upon me.” (Treatise: 1.3.VIII, 83)

This theory of judgement is significant since it determines the criteria which a successful explanation of a judgement must meet. In particular, since a judgement or belief is simply a very lively and vivid idea, an adequate tracing of the genesis of such a judgement must determine:

1. How is the relevant idea acquired?
2. What confers the requisite FLV rendering the idea a judgement?

---

<sup>17</sup> This conception of belief/judgement has been subject to a barrage of criticism from commentators (e.g. Bennett, 1966: 54-5; Stroud, 1977:70-72), although Zangwill (1998) points out that the alternatives appealed to by critics are problematic also.

#### 1.4. A worry about the theory of ideas

There can be little doubt that advances in philosophy and science over the past two hundred and fifty years have left Hume's account of the mental furnishings of moral subjects looking desperately simplistic. Such a fact is neither surprising nor, it seems to me, especially interesting. What is interesting however, is the validity of the general theoretical framework of which Hume's specific proposals are an example: the theory of ideas. Specifically, is it legitimate, given naturalistic aims and method, to explain the genesis of judgements in terms of a succession of mental particulars and relations holding of them? The answer generally given to this question in the literature is *no*. In fact, it has been common for commentators to see a deep inconsonance between Hume's science of man on the one hand, and his theory of ideas on the other. It is widely asserted that Hume's theory of ideas and his naturalism are incompatible in the sense that the former distorts, warps, and is generally inimical to naturalistic inquiry. The theory of ideas is a fetter upon Hume's naturalism, something which a purely naturalistic approach would be free of.<sup>18</sup>

“One thing that works against a consistent and comprehensive naturalism in Hume's own thought is his unshakable attachment to the Theory of Ideas.” (Stroud, 1977: 224)

Call this the inconsistency thesis (IT). In this section and the following section, I shall address, in turn, the two main arguments which have been proposed in favour of IT and aim to refute them both. Far from being inimical to his naturalism, I think that Hume has good reasons for adopting the theory of ideas and utilising it as the theoretical framework within which the genesis of judgements is framed. Specifically, the theory of ideas can and should be interpreted simply as a thesis about how human beings are to be understood if they are to be understood at all and as such, it is *integral* to Hume's science of man.

The first argument which has been advanced in favour of IT concerns the nature or metaphysics of ideas. Specifically, some commentators have thought that the *mentalistic* character of theory of ideas renders it naturalistically unpalatable. For example, Howard

---

<sup>18</sup> Stroud, (1977, 2006); Mounce, (1999); Rupert Read (2000: 170) talks sweepingly of Hume's “(on almost any reading we are used to) ‘unfortunate’ methodology centred around his so-called ‘theory of ideas’”. The only philosopher I am aware of who has actively argued the contrary is Fodor (2003).

Mounce, in an otherwise positive assessment of Hume's science of man, laments the "mentalism" which the theory of ideas imposes upon the inquiry: "Hume treats an idea as a kind of object, differing only in that it occurs in the mind." This is said to be a "difficulty", the problem apparently being that "the empiricist reduces the mental to what is in the mind... In short, he dissociates it from the public realm and associates it with the private" (Mounce, 1999: 29)

According to Mounce, a naturalistic account of man will not admit of a murky, unobservable realm of mental particulars in 'peoples' heads', but will replace these with publically observable phenomena. Specifically, rather than thinking of an idea as a private image before the mind's eye, ideas should be understood as capacities to *do* certain things in certain circumstances. What exactly these capacities are is open to debate. One possibility is that they are recognitional in character. In this case, to possess the idea of a fox is to possess the ability to distinguish or sort foxes from non-foxes. Another possibility is that to possess an idea is to be able to draw certain inferences e.g. if x is a fox, x is a mammal etc. However this 'doing' is construed, the crucial tenet is that idea possession is to be identified with the ability to function or act in certain ways in certain contexts: that is, it is a dispositional property. As Pears (1990: 25) puts it:

"When a concept manifests itself as a particular image occurring in a person's mind at a particular time, it cannot be just identified with that image... [Rather, we] have to add that it is only the image with its special function."

Equally, Stroud (1977: 227) insists that we do not have any clear sense of what an idea is unless it is identified with a unique "function, or point" (see also Flew, 1986: 25). The significance of this functionalist proposal can be sharpened by contrasting it with an alternate and, *prima facie*, appealing way to understand the relationship between the possession of an idea and the performance of a certain function. For example, one might think that just as a lumberjack is able to cut down a tree due to his possession of a chainsaw, so too one is able to perform a function, such as sorting foxes from non-foxes, precisely because one possesses the requisite idea (for instance, a mental image of some sort). In this case, the possession of the idea enables a function, but is no more *constituted* by that function than the lumberjack's possession of a chainsaw is constituted by his ability to cut down trees.

It is precisely this picture which Mounce and others find pernicious and want to replace. According to their functionalist proposal, idea possession should be likened to



knowing how to ride a bike. An agent knows how to ride a bike if they can perform certain acts: can peddle without losing balance, can accelerate and brake in appropriate situations etc. To be able to perform in these ways *simply is* to know how to ride a bike; there is no extra question of ‘knowledge possession’ to be answered over and above the ability to so act; the function constitutes the possession. Similarly for ideas, the functionalists argue that to possess an idea *simply is* to be able to act or function in certain ways. Possession and function are not merely conditionally connected, but identified. This identification forces ideas from the private confines of the mind and into the bright light of the public domain. If to possess an idea is to be able to do certain things, idea possession can be subject to third party observation and verification.

This functionalist proposal is persistently recommended as a way to save Hume’s science of man from the perfidious spectre of mental particulars. The underlying premise is that if these theoretical constructs are not, at some point, traded in for an ‘act-based’ or ‘function-based’ coinage then they are naturalistically unacceptable. However, why should we accept this premise? Despite the popularity of the functionalist alternative in the literature, it is hard to find explicit discussion of why or how Hume violates his naturalism by postulating mental particulars. For example, Flew (1986: 25) suggests that Hume never fully recognised that “the keys [to understanding] lie in abilities and dispositions to speak and behave”, persisting instead with mental imagery. However, Flew offers no explanation of why the keys lie in abilities and dispositions. Similarly, Mounce says that:

“On the empiricist view, an idea is an image, which occurs ‘in the mind’. Only later do we give it outward expression in language. The process evidently runs the other way. Certainly I can think to myself, without expressing my thoughts; but then I have already learned to express myself in language.” (Mounce, 1999: 29)

Since ‘language’ here denotes something like the ability to engage in a type of publically observable behaviour i.e. produce utterances or marks in appropriate circumstances, this is a (linguistic) formulation of the functionalist thesis. Yet, as far as I can see, there is nothing “evident” about this at all and Mounce makes only the briefest of attempts to support this assertion. He argues that even though it is possible to read silently, reading aloud is the “primary activity” since this is how everyone learns to read. Similar points are said to apply in the case of ideas. However, in what sense reading aloud is supposed to be ‘primary’, and

why, besides, it is a suitable analogy for ideas (many people cannot read, all people form ideas), he does not comment upon.

I think that this functionalist critique has been underpinned, at least partly, by a mistaken conception of what Hume's naturalism amounts to. Specifically, it is conflated with a naturalism of a more contemporary philosophical hue. Nowadays, naturalism in, say, the philosophy of mind or language typically involves a commitment to the ontological primacy of a certain scientific vocabulary, usually physics i.e. it is closely tied to some form of Physicalism (Cabrera, 2011: 46-47). The great challenge for the philosophical naturalist is whether our ordinary thought and discourse can be reconciled with this ontological commitment. That is, people talk and think, or at least appear to talk and think, about things like value, modality, meaning, mental properties. The defining problematic of *this* naturalism is whether these domains of speech, thought and practice can be accommodated or, as Huw Price puts it, 'placed', in a world which is solely comprised of a physical ontology (Price, 2004).<sup>19</sup> How can these vast tracts of human experience be shown to be legitimate without relying upon an implausible ontology of entities? Or does a 'naturalistic' ontology render much of our day to day practice and common sense thought erroneous?

Now, if this were the type of naturalism inherent to Hume's science of man, there may be good reason to suppose that his postulation of mental particulars is unacceptable on ontological grounds and should be replaced with descriptions of behaviour or dispositions to behave.<sup>20</sup> However, the difference between Hume's science of man and this type of naturalism should be evident. For one thing, Hume's eschewal of epistemic concerns means that he is quite immune to this 'placement' challenge. He is not questioning how our everyday talk and practices involving, for example, causal necessity can be 'placed' in the world (see, for example, *Treatise*: 1.3.XIV). More crucially, there is no trace of an ontological commitment in the *Treatise*. Indeed, in positing laws of association Hume was tacitly repudiating the Physicalism of the day which was founded upon the idea of contact mechanics (Hinzen, 2006: 6-9). Just as Newton had reintroduced 'occult' forces to explain the behaviour of bodies in nature, Hume was pursuing analogous "secret springs and principles" of the mind which determine the behaviour peculiar to moral subjects (*Enquiry*: 1.15, 93). At best, Hume was agnostic about whether these secret springs could be reduced to

---

<sup>19</sup> To be clear, although the terminology is Price's, it is not a problem he endorses or accepts.

<sup>20</sup> Or indeed, something else altogether e.g. neural activity.

a more primitive scientific vocabulary. He says that they are “original qualities of human nature, which I pretend not to explain” (Treatise: 1.1.IV, 15).

Hatfield (1990: 16-17) demarcates the naturalism of Hume’s science of man by labelling it ‘methodological naturalism’, and contrasts it with the ‘metaphysical naturalism’ of the contemporary philosophical scene.<sup>21</sup> For Hume, something is naturalistic if it is motivated by the naturalistic method: one of careful observation and experimentation. That is, what is natural is what features in our best scientific theories. It follows that if Hume’s theory of ideas is to be shown to be inconsistent with his science of man, it is naturalism of the methodological kind, not the metaphysical which must be targeted.<sup>22</sup> Yet, it is clear that perceptions do explanatory work for Hume. For example, Hume was keenly aware that we are capable of a vast number of thoughts, many of which, maybe most of which, we have had no experience of e.g. golden mountains, spiteful giraffes. He accounted for this by saying that while our experience is indeed limited and fixed, it supplies us with atomistic mental elements with which we can construct the most bewildering array of thoughts. Similarly, our stream of thought adheres to a certain order or regularity which is a product of our experience. If my mother always used to feed me pomegranate and apple slices for dessert, the thought of one will naturally prompt the thought of the other. This is surely not because there is special link between pomegranates and apple slices, but only a link between my *ideas* of them. Furthermore, according to Hume, a framework of this sort enables one to make substantive insights into the origins of our beliefs (some examples of which I shall discuss in Chapter two). Ultimately, the plausibility of the theoretical framework Hume adopts “rests upon the success of his explanatory program” (Flage, 1990: 36).<sup>23</sup>

Given this, Hume’s postulation of mental particulars can be understood in abductive terms, as an inference to the best explanation of certain facts about human beings. Insofar as they play an explanatory role in a science of man, they are as natural as any other unobservable entity postulated by a scientific theory. More generally, this aspect of the theory of ideas is just a claim about how human beings are to be understood if they are to be understood at all, and consequently, cannot be used as an argument in favour of IT. As Fodor

---

<sup>21</sup> Alternatively, adopting Price’s (2004) distinction of ‘subject naturalism’ and object naturalism’, we can say that Hume’s science of man is subject naturalist but not object naturalist (a selection Price himself recommends).

<sup>22</sup> From here on in, ‘naturalism’ will refer to *methodological naturalism*, except where otherwise specified.

<sup>23</sup> These explanatory considerations bear more than a passing resemblance to many of those appealed to by Fodor in support of his Representational Theory of mind, according to which thought involves the manipulation of mental particulars. For example, in order to account for the productivity of thought (pace Hume), Fodor (2003: 135) argues that it is necessary to postulate a realm of mental particulars which compose.

(2003: 8) puts it, “Hume saw that accepting... the “Theory of Ideas” is central to constructing an empirically adequate account of cognition”.

Two clarifications are in order. Firstly, none of this entails that the theory of ideas *really is* central to an empirically adequate account of cognition. Hume, not to mention Fodor, might just be wrong about this; perhaps an empirically adequate account of cognition need not make reference to mental particulars, but can rely instead of the kind of dispositional traits so commonly recommended by Hume commentators. However, it should be clear that this amounts to no more than the virtual truism that any scientific thesis is open to doubt, is never epistemically necessary, and is always at risk of being displaced by alternate theories; none of this renders a theory *non-naturalistic*, and nor can it do so with respect to the theory of ideas. Secondly, I stated at the outset of this section that I wanted to defend not the details of Hume’s theory of ideas, but the consistency of this *kind* of theoretical framework with his naturalism. Crucially, nothing I have said entails or rests upon the legitimacy of Hume’s specific conception of ideas as *images*.<sup>24</sup>

### ***1.5. Another worry about the theory of ideas***

Even if Hume’s metaphysics of ideas is naturalistically acceptable, this does not address the most prominent objection to the theory of ideas. This objection concerns the empiricist model of perception enshrined in the theory. According to the theory of ideas, being in a mental state is to have a particular mental atom or arrays of atoms before the ‘mind’s eye’ and this includes perceptual states. On this view, to be in such a state, it is both a necessary and a sufficient condition that one have a certain kind of mental object, a lively and vivid image or *impression of sensation*, ‘in mind’. As such, Hume’s model of perception is a type of sense

---

<sup>24</sup> There are a whole cluster of related problems for this view. One especially pertinent problem concerns the possibility of abstract ideas. Roughly, images seem too fine grained to allow for abstract ideas. For example, it is hard to see how the thought ‘A killed B’ could be a complex image, because an image would necessarily depict the manner of killing. What would a thought ‘A killed B’, unspecified for manner of killing e.g. stabbing, shooting, etc., *look like* (Jackendoff, 1996: 10-11)? Recognising this problem, Hume does propose a theory of abstract ideas which is clearly aimed at addressing this difficulty. Hume’s theory of abstract ideas states (a) there is no idea which is *truly* abstract, i.e. unspecified with respect to certain properties of the idea, e.g., the colour of a cat, since “’tis utterly impossible to conceive any quantity or quality, without forming a precise notion of its degrees”, and (b) instead, there are resemblance classes of particular ideas which, linked by a common name, are ‘revived’, though not all actually present before the mind, by the utterance of the term (Treatise: 1.1.VII, 19). Even setting aside the difficulty regarding abstract ideas however, there are numerous other problems with the view that ideas are images, such as how to explain logically complex representations, or the representation of numbers.

datum theory. It is widely agreed that Hume accepts this model unquestioningly, as a foundational premise of his inquiry:

“...Hume never seriously considers any evidence for or against the theory of ideas”  
(Stroud, 1977: 111; see also, Flew, 1986: 14-29)

In the context of IT, the crucial issue is not merely whether this model is plausible or implausible, but whether it is inconsistent with a naturalistic inquiry. In other words, does it affect or determine, in some significant sense, the contours of Hume’s naturalism? Such a determination is suggested by Stroud’s comment that in Hume’s science of man,

“Flesh and blood human beings as they walk the earth do not come into the story at all. Nor does the earth for that matter, or any of the objects that naturally exist in it.”  
(Stroud, 2006: 346)

Hume goes to great lengths detailing the character and properties of our perceptions. But of non-perceptual objects, events and properties, Hume has almost nothing to say. Hume talks of impressions of chairs, but not of chairs; of impressions of apples, but not of apples. For example, nearly a third of way the way into the *Treatise* Hume springs the question, ‘why do we attribute to objects an existence distinct from our mind and perception?’ This is despite spending the previous hundred pages freely discussing our impressions of letters, fires and billiard balls. Similarly, all positive discussion of relations is limited to those which hold only between perceptions. Ideas are copied from impressions; ideas can be compared by means of resemblance, identity, space and time, quantity, quality, contrariety and causality (*Treatise*: 1.1.V). Yet, of relations between impressions and non-perceptual objects e.g. a causal relation, Hume is strikingly silent. It is hard to overstate the significance this eschewal of non-perceptual objects has upon the overall character of Hume’s naturalistic enterprise. Hume is interested in a “descriptive – and explanatory – science” (Biro, 1993: 38), but the only thing which does any describing or explaining are our own perceptions. Our judgements are explained by tracing their genesis back to our ideas and their association. In turn, ideas are copied from impressions. Yet, here the explanatory buck stops. Impressions comprise a primitive level of theoretical analysis for Hume. Call this Hume’s constraint (HC).

“Nothing in nature is called on to explain our beliefs. Or rather... nothing beyond momentary and fleeting perceptions and those principles of the imagination that hold in the mental realm of the comings and goings of perceptions.” (Stroud, 2006: 345)

HC seems to deprive Hume of the resources which an adequate science of moral subjects surely requires. For example, my current behaviour presumably has something to do with the fact that there is a desk with a computer on it in front of me. Yet, all Hume could say is that I have the complex impression of a desk with a computer atop it. Surely a more complete naturalistic account would make reference to the *desk itself* and achieve greater explanatory depth thereby!

This constraint upon Hume's naturalism is attributed to his theory of ideas, or more specifically, the model of perception which derives from it. To say that perception is simply a matter of mental impressions being 'before the mind's eye', is to say, so goes the objection, that moral subjects are only ever directed or non-inferentially acquainted with their own perceptions. Knowledge of anything non-perceptual can only be achieved via inference from these perceptions.<sup>25</sup> However, once non-inferential epistemic access is restricted to our own perceptions, we can never get in 'behind' or 'underneath' them to make direct contact with a non-mental reality. According to this picture, instead of our perceptions affording us an opening unto the world, they screen our thought from it. This is, of course, the famous veil of appearance. According to this objection then, call it the *epistemic objection*, it is Hume's theory of ideas which limits Hume's naturalism to the perceptual domain. The theory comes at the expense of the objects we ordinarily take ourselves to encounter in experience: the kind of objects which, *prima facie*, we would expect a science of man to appeal to. HC is simply a consequence of Hume's model of perception which is, in turn, an unjustified presumption on Hume's part. As Stroud (2006: 345) puts it, theory of ideas is a "troubling constraint upon Hume's naturalism. It is not a constraint upon naturalism as such, or naturalism in every form, but on Hume's naturalism in particular." A complete naturalistic program would make reference to such objects and therefore a complete naturalistic program is inconsistent with the theory of ideas. Indeed, this was essentially how Reid saw the matter. According to Reid, the theory of ideas is a poisonous element which inevitably corrupts Hume's conclusions: "it [the theory of ideas] had a specious appearance of both innocence and beauty; but ... carried in its belly death and destruction" (Reid, 1764: 75-76). Reid's proposed solution was to dispense with the theory of ideas and pursue a naturalistic approach which, most notably, appealed to the everyday objects of common sense (cf. Norton, 1982: 192-200; Williams, 2004: 295).

---

<sup>25</sup> Hume is often said to have taken the empiricist views of his predecessors to their logical and sceptical conclusions by arguing that there is no epistemic grounds for moving from our perceptions to what is non-perceptual.

I think that this objection completely misconstrues the relationship between Hume's model of perception and his naturalism. I will begin my argument with reasons to think that the relation, as portrayed by the epistemic objection, is mistaken. Subsequently, I shall develop an alternate interpretation. The point of departure for this argument is some comments Hume makes in the *Treatise*. Although, as I have noted, Hume does not engage in any positive discussion of non-perceptual objects, he does make some 'negative' remarks which are very important:

“Tis certain, that the mind, in its perceptions, must begin somewhere, and that since the impressions precede their correspondent ideas, there must be some impressions, which without any introduction make their appearance in the soul. As these depend upon natural and physical causes, the examination of them wou'd lead me too far from my present subject, into the sciences of anatomy and natural philosophy.”  
(*Treatise*: 2.1.I, 213)

Furthermore, he elsewhere reinforces this theme, suggesting that questions regarding the origins of impressions or sensations are properly left to “anatomists and natural philosophers” (*Treatise*: 1.1.II, 12). These statements are highly indicative. Firstly, they are incompatible with the idea that Hume's model of perception constrains naturalistic inquiry in the sense of foreclosing epistemic access to non-perceptual objects. If this were really how Hume understood the ramifications of this model, we would not expect him to advise us that anatomists can investigate the causes of our impressions. After all, if this model of perception precludes reference to the causes of our perceptions for the moral philosopher, it does so for the anatomist also.

Secondly, Hume expresses in these remarks a very interesting attitude towards the topic of non-perceptual objects: the topic is portrayed as being essentially orthogonal to the kind of inquiry he is undertaking: “the examination of them wou'd lead me too far from my present subject”. As Marjorie Grene (1994: 164) puts the matter:

“There is surely a real world we are living in, and are a part of. Only we are treating that world here in terms of moral rather than natural philosophy; we are starting with our perceptions, not with the anatomy or physiology that would attempt to describe

their causes... the bodily events, movements of the spirits and the like, that presumably cause our perceptions can indeed be studied - elsewhere and by others.”

The suggestion here is that limiting inquiry to the mental domain of perceptions is not *imposed* upon a science of man, but is something which follows or derives from the character of the inquiry itself. Scientists of man concern themselves solely with perceptual objects, because this is the appropriate subject matter for such an inquiry, in the same way that other sciences have their own distinct subject matter. It follows from this that what I earlier called ‘Hume’s constraint’, is not really a *constraint* at all, at least not in the sense of being externally dictated. Rather, Hume portrays it as an ‘internal’ feature of the kind of naturalistic inquiry he is engaged in. But then, evidently, if there is no constraint, Hume’s model of perception cannot be its source.

However, if Hume’s model of perception does not constrain his naturalism, what is the relationship between this model and his naturalism? Consider an observation Fodor makes about empiricist or sense data theories of perception. Fodor points out (1975: 42-53) that such theories have, historically, generally performed a “double duty”. On the one hand they are meant as accounts of how beliefs can or ought to be justified. They state what it is subjects have direct or secure epistemic access to, and then show how our beliefs can (or cannot) be built upon such a foundation. On the other hand however, empiricist models of perception have also acted as *psychological* accounts of perception: purporting to depict what is involved in perceptual processing as a matter of fact. Now, evidently, these two duties are related; how one is met is likely to have implications for the other. *Nevertheless*, there is question as to which of these duties is dominant or primary and this is particularly important with respect to Hume. Commentators have almost universally assumed that Hume’s model of perception principally serves the epistemic duty, virtually equating it with the veil of appearance. If it is understood in this way, as being, first and foremost, about what we do and do not have epistemic access to, then it is easy to see how this constrains naturalistic inquiry in the way Hume scholars have claimed. However, commentators have neglected the possibility that it should be understood as principally answering to Fodor’s psychological duty. In this case, far from being inconsistent with a naturalistic approach, the model would be something which *emerges* from this approach and is premised upon it. Rather than being an unquestioned presupposition, it would be something which naturalistic considerations lead him to, answering to whatever standards of empirical adequacy are relevant. What I want to



claim is that this is how Hume's model of perception should be interpreted: an attempt to provide a psychologically accurate account of what is involved in perception. Although such a view certainly raises epistemic questions (see fn. 31), these are *secondary*. That is, they are consequences of an account justified on psychological grounds.

It is worth underlining from the outset the appeal and attraction of pursuing this interpretation of Hume's model of perception. Even setting aside the noted incompatibility of the epistemic reading with many of Hume's remarks, a psychological reading promises to be a far more *charitable* option in at least two respects. Firstly, if the model is read as an epistemic claim about what we do and do not have cognitive access to, then Stroud is correct that Hume never considers any evidence for it. Secondly, on the epistemic reading, the model is inconsistent with Hume's overarching philosophical program: his science of man. Now clearly, taken alone, such considerations do not prove anything: it may just be that we have to attribute these philosophical aberrations to Hume. *Nevertheless*, these considerations do justify an exegetical preference for an alternative interpretation which does not indict Hume in these ways. This is what the psychological interpretation promises to be and insofar as it can be coherently developed it should be preferred in any charitable reconstruction of Hume's science of man.<sup>26</sup> To this interpretation I now turn.

One of the most striking facts about perceptual or experiential states, and one which is especially important from the perspective of a science of man, is that they play a causal role in the behaviour of moral subjects, including of course their *judgements*. For example, consider how people make judgements about unobserved matters of fact on the basis of what is observed, a phenomenon which Hume is uncommonly interested in (see Section 2.1). If I peer out of a window and observe that the street is covered in puddles, I may well judge that it has been raining. Now, what is significant about this is not merely that perceptual states play a causal role in producing judgements, but that what is causally efficacious is the specific character or *content* of a perceptual state. If I peer out of a window and observe a thick layer of snow covering the road, I would presumably be led to quite a different

---

<sup>26</sup> It might be quite reasonably pressed whether I am offering an exegetical interpretation of Hume's intentions, or whether the argument is really a philosophical one, i.e. that there is a way to reconcile Hume's theory of ideas and his naturalism, irrespective of whether he in fact saw it this way. The idea of a 'charitable reconstruction' is ambiguous in this respect. On the whole, I am primarily interested in offering a philosophical argument rather than an exegetical account. However, this account is motivated by remarks Hume himself makes. Consequently, it might be accurate to say that the argument here is inspired by Hume, even if ultimately he would not have endorsed it (after all, in order to establish this latter thesis, a much wider consideration of his related remarks would be required – something I shall not be attempting).

judgement. In other words, the judgements moral subjects produce depend upon the content of their perceptual states. Consequently, we are immediately confronted by the challenge of stating what determines the content of perceptual states, such that different content can play different causal roles. This is a challenge for any science of man which purports to explain the behaviour of moral subjects.

In Section 1.4.II of the *Treatise* Hume discusses certain ‘experiments’ and I think that these have a crucial bearing upon how perceptual content is to be understood. Each of these experiments indicate the dependence of perceptual content upon the dispositions of our organs. For instance, when we press one eye with a finger, we immediately perceive all objects to become double. Similarly, qualities such as colour and figure change also in line with our “sickness and distempers”. These and an “infinite number of experiments of the same kind”, prove “that our perceptions are dependent upon on our organs, and the disposition of our nerves and animas spirits” (*Treatise*: 1.4.II, 162-163). Essentially, Hume is appealing to the kind of illusion and hallucination considerations which are commonplace in philosophy of perception literature.

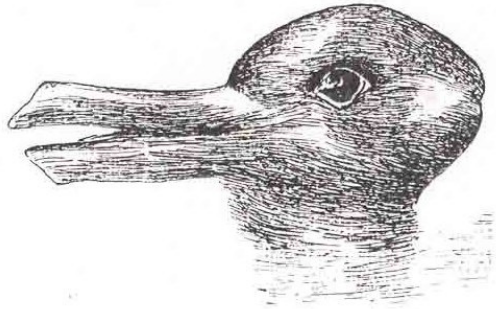
These experiments have generally failed to impress commentators. Stroud immediately dismisses them on the grounds that “it of course does not follow from this that what we perceive is a ‘momentary’, ‘fleeting’, ‘internal and perishing’ thing which depends for its existence on the mind that perceives it” (Stroud 1977: 111; cf. Flew, 1986: 33). However, while Stroud is right to dismiss the idea that these experiments establish that we are only ever directed acquainted – where *acquaintance* is an epistemically loaded concept - with our impressions, i.e. the idea that they establish a veil of appearance, these experiments can, I think, reasonably be taken to indicate a different conclusion: that the content of our perceptual states is not fixed or individuated by features of the physical or social environment.<sup>27</sup> Roughly, these experiments show that features of the external environment can be held constant while perceptual content varies, as in illusions, and even that perceptual content can remain constant despite variance in the external environment e.g., as in hallucinations. In other words, there is no deep individuation relation holding between perceptual content and nature of a subject’s environment.

---

<sup>27</sup> That is, I am essentially attributing to Hume the view that perceptual content is narrowly individuated (cf. Kriegel, 2008).

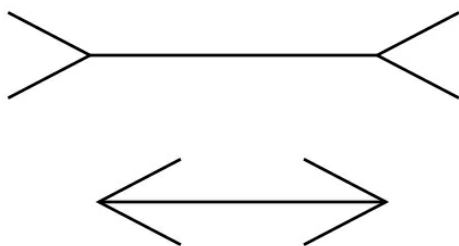
There is a sense in which there is nothing particularly revelatory about this; a similar thesis is widely acknowledged by cognitive scientists: “all physical events are intrinsically ambiguous, in the sense that they are subject to various interpretations” (Pylyshyn, 1980: 112).<sup>28</sup> Consider two simple, but illustrative, examples. Firstly, Wittgenstein’s well known rabbit-duck illusion.

Figure 1.



The image can be perceived as a rabbit or a duck, but not both at the same time and as such, it is associated with two incompatible perceptual contents. It follows from this that even for such a relatively impoverished visual experience, the relevant stimulus, or ‘facts about the physical and social environment’, cannot fix perceptual content. Since the stimulus, can remain constant while the perceptual content of the subject varies, the latter cannot be individuated on the basis of the former. Alternatively, consider the Muller-Lyer illusion.

Figure 2.



Subjects perceive the top line to be longer than the bottom line, despite their being equivalent in length. Now, evidently there is no *ambiguity* here as there is in the duck-rabbit case. But the example serves to generalise the relevant notion of ambiguity: namely, as that according to which the nature of the relevant environmental stimulus underdetermines or is too coarse

<sup>28</sup>This comparison is especially worthy of consideration given the parallels which have been drawn between Hume’s science of man and cognitive science e.g. Biro (1993), Fodor (2003), Garrett (1997:40). In Section 3.3, I discuss how this underdetermination thesis is a standard assumption in moral psychology.

grained to individuate the content of a subject's perceptual state. This is a completely general phenomenon. As Fodor puts it:

“...if you want to know how the organism will respond to an environmental event, you must first find out what properties it takes the event to have” (Fodor, 1975: 55)

There is no property of the stimulus in the Muller-Lyer case which determines that the lines are perceived as they are. But then, it follows that a subject's perceptual content, while *elicited* by the visual stimulus, cannot be fixed or specified by it; the content is underdetermined. For this reason, Fodor (1975: 55) draws a principled distinction between ‘distal stimulus’, whatever features of the environment may elicit a perceptual state, and ‘proximal stimulus’, which determines the content of such a state.<sup>29</sup> It seems to me that it is exactly this distinction which Hume is appealing to when he says:

“The table which we see seems to diminish as we move further away from it; but the real table, which exists independent of us, suffers no alteration.” (Enquiry: 12.9, 201)

So the content of a perceptual state is not fixed or individuated by objects in the environment. But then what is it fixed by? What is Fodor's ‘proximal stimulus’? The natural alternative here is that it is individuated on the basis of the intrinsic properties of the subject. This is essentially what Hume concludes, postulating a domain of particulars in the mind of every subject, *impressions of sensation*, which fix the content of a perceptual state. These mental atoms have their content essentially,<sup>30</sup> and serve as the foundation for all other mental content. Although impressions are not publically observable, they are available to introspection.

“For my part, when I enter most intimately into what I call *myself*, I always stumble upon some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure.” (Treatise: 1.4.VI, 193)

---

<sup>29</sup> This thesis is, for example, standard in moral psychology (see Section 3.3). Similarly, consider Chomsky's claim (2000: 23) that “in the study of determination of structure from motion, it is immaterial whether the external event is successive arrays of flashes on a tachistoscope that yield the visual experience of a cube rotating in space, or an actual rotating cube, or stimulation of the retina, or optic nerve, or visual cortex. In any case, “the computational investigation” concerns the nature of the internal representations used by the visual system and the processes by which they are derived.”

<sup>30</sup> Although the idea of mental atoms having content essentially may sound tendentious, it is not clear to me that it is anymore tendentious than saying that propositions have truth conditions essentially – a widely held and apparently respectable philosophical thesis (though for a critique, see King, 2009).

Moreover, and more importantly, impressions can be postulated in the minds of every person, despite being private, insofar as they play an explanatory role in our best account(s) of their behaviour. That is, the characterisation of perceptual states in terms of array of mental impressions resembles the kind of ‘inference to the best explanation’ that I argued in the previous section Hume’s account of ideas should be understood as. Hume’s model of perception is, first and foremost, a psychological thesis, not an epistemic one about what we do and do not have cognitive access to.<sup>31</sup> It is an attempt to answer a question central for a science of man: what determines perceptual content such that it plays a causal role in the judgements of moral subjects? As such, the model is justified insofar as it is necessary for (or plays a role in) an empirically adequate science of man. Furthermore, it is worth noting that this line of reasoning and conclusion is by no means alien by contemporary lights:

“I have been saying sometimes that attention is drawn by an object in the world, sometimes that it is drawn by a percept in the head. Which is right? If we want to be precise, we have to say a percept. Attention is a process going on in the brain - so it *cannot* be directed at things moving out there in the external world.” (Jackendoff, 1996: 22)

*Crucially*, nothing about this position precludes there being non-perceptual stimuli which elicit our impressions. In the case of the diminishing table, Hume explicitly contrasts the image present to the mind with the *real table*. In fact, as far as I can see, the view is consistent with the claim that any impression of sensation *must* have a non-perceptual cause(s) (Fodor’s distinction between proximal and distal stimulus presupposes as much). The point is rather that since these causes do not fix the content of a perceptual state, they are not the appropriate subject matter for a science of man. Since, *ex hypothesi*, it is the array of impressions moral subjects find before their mind’s eye which fixes perceptual content, the scientist of man is justified in ‘starting’ with these.

---

<sup>31</sup> This is not to say that this model does not raise interesting epistemological questions. The classic problem with individuating content non-relationally, solely on the basis of the intrinsic properties of the individual, is that it does not guarantee that said content puts subjects in ‘cognitive contact’ with the world (see Kriegel, 2008). Indeed, Hume persistently warns us that there is no good reason to suppose that objects are as we take them to be. Nevertheless, these considerations can be separated from the core justificatory basis of the model. As Norton (1993: 8) puts it: “Hume gave the “way of ideas” a kind of phenomenological turn. That is, his primary concern... is with our perceptions, qua perceptions, with perceptions as, simply, the elements or objects of the mind and not as representations of external existences”. In other words, Hume was primarily interested in the way in the causative properties of perceptual content can be understood, and only secondarily interested in the way in which, if at all, such content puts us in cognitive contact with the world.

“Hume has declared at the outset that our perceptions are caused by things or events other than themselves; as moral philosophers we are just not asking about this relation.” (Greene, 1994: 171)

By way of summation for this rather lengthy section, consider what I have sought to establish and what I have *not* sought to establish. Taking the positive component first, I have argued that Hume’s theory of ideas, more particularly, the model of perception it entails, is consistent with Hume’s naturalism, contrary to what is widely claimed. Breaking with the standard interpretation of this model, I have argued that there is good reason to interpret it as a psychological account of perception, one which addresses a question central for a science of man: what is perceptual content such that it plays a causal role in the judgements of moral subjects? In this sense, whatever problems may arise for Hume’s model of perception (see Section 2.3), it is not only perfectly naturalistic, but integral to his naturalistic program.

## Chapter 2. Applying Hume's Naturalistic Approach

---

Hume believed that his theory of ideas enabled him to draw the kind of conclusions a science of man aims at: why moral subjects form certain judgements. In this chapter, I shall examine in turn two cases Hume discusses: causal judgements and judgements of the existence of external bodies. In the final section, I shall argue that inadequacies with his account of our judgements of external bodies are indicative of a very serious shortcoming for his science of man, one which strikes at the heart of how judgements are conceived of and treated in this tradition.

### 2.1. Causation

There is perhaps no topic, says Hume, which has prompted more disputes amongst both ancient and modern philosophers than that of the efficacy of causation, and significant portions of both the *Treatise* and the *Enquiry* are devoted to it. That moral subjects judge objects to be causally related is a trait of fundamental importance to their day to day reasoning and behaviour. In particular, all of our beliefs regarding unobserved matters of fact rely upon causal relations.

“If you were to ask a man, why he believes any matter of fact, which is absent; for instance, that his friend is in the country or in France; he would give you a reason; and this reason would be some other fact; as a letter received from him, or the knowledge of his former resolutions and promises.... it is constantly supposed, that there is a connexion between the present fact and that which is inferred from it. Were there nothing to bind them together, the inference would be entirely precarious.” (Enquiry: 4.4, 109)

So, the causal judgements Hume is interested in are those whereby, on the basis of something observed, we come to believe an unobserved matter of fact. If I observe the sun low on the horizon, I am moved to believe that the surrounding buildings are casting shadows. Or if I observe that the bread bin is empty, and remember that it had contained a loaf of bread, I am led to believe that someone ate the bread. These transitions from what is observed to belief in

that which is unobserved are “founded on the relation of *Cause and Effect*” (Enquiry: 4.14, 113). *That* people make causal judgements of this form is the datum Hume wants to explain. “We must enquire how we *arrive at the knowledge* of cause and effect” (Enquiry: 4.5, 109) [my italics]. Although causation was an old philosophical topic, debate had been speculative or metaphysical in character: about the existence of causal relations holding of objects irrespective of how they are conceived. What is, in many ways, the more modest investigation into the origins of our causal judgements had been neglected and it is this Hume wants to redress: “this is what I principally find wanting in their reasoning, and what I shall here endeavour to supply” (Treatise: 1.3.XIV, 123). He is concerned with “the cause of our causal beliefs” (Flage, 1990: 92)

Hume begins his investigation by examining two things which would be considered causally related and asking what impressions we receive when we experience them. The aim here is to uncover some perceptual quality from which we copy the idea of causation. For any two objects said to be causally related, we find that there are certain relations which hold of them. Firstly, the objects are contiguous in time and space. Admittedly, on occasion, a certain distance may separate cause and effect, but here we quickly find that they are linked by a chain of causes which are themselves contiguous. For example, consider ‘cause to die’ debates where an action causes a patient’s death, though the death occurs at a later time. Despite the temporal distance, we find a chain of contiguous causes linking the original action to the death. Secondly, temporal contiguity is always of a certain order. Cause and effect cannot be simultaneous and the order is always from cause to effect.

This, however, exhausts our inspection of impressions and it is one which fails to yield the sought for result. Objects can be contiguously and successively related to another without being considered cause and effect. If a fly lands on my computer screen, and then my screen saver starts, I do not consider them causally related, despite the contiguity and order of the two events. Causation entails that there is a *necessary connection* also. By ‘necessary connection’, Hume means that an object causes another only if it can be said to produce it, or that without the first object, the second would not have appeared. ‘Necessity’ is synonymous with ‘power’ or ‘production’. It is this quality which appears to be absent from our impressions.

Here, there is something of an interlude in Hume’s thought, and it is a significant one. The copy principle states that all our simple ideas are copies of simple impressions. Yet, at



least after an initial inspection, there seems to be no impression corresponding to our idea of necessity. This should suggest two possibilities:

1. The copy principle is false (in which case, thoroughgoing empiricism is untenable)
2. Humans do not have an idea of necessity.

Despite considering the copy principle to be the ‘first principle of the science of human nature’ and too well established to be dispensed with, Hume still comes closer to accepting (1) than (2):

“Shall the despair of success make me assert, that I am here possest of an idea, which is not preceded by any similar impression?” (Treatise: 1.3.II, 64)

Hume’s refusal to consider the possibility that we have no idea of necessity is instructive. It reinforces the explanatory, as opposed to epistemic, aims of the investigation. After all, if the question was whether our causal judgements are *justified*, having inspected our impressions and found no grounds for the idea of necessity, we might expect Hume to contemplate the possibility that our causal beliefs are not justified precisely because we lack the requisite idea of necessity. The reason Hume does not contemplate this possibility – even though the alternative may be that he needs to abandon the copy principle – is that *that* we have the idea of necessity is a presupposition of his inquiry. Our causal judgements are the explananda for Hume, and insofar as a causal judgement presupposes an idea of necessity, to admit that we have no such idea would be to conclude that a naturalistic study of our causal judgements is misguided or mistaken. That is why Hume never doubts that we have the idea of necessity; he is aiming to explain our causal judgements by tracing their genesis, not establish their epistemic merit.<sup>32</sup>

Having failed to find an impression corresponding to our idea of necessity, Hume adopts a more indirect approach, pursuing the question: “why we conclude, that such particular causes must necessarily have such particular effects, and why we form an inference

---

<sup>32</sup> Although this point may seem obvious, at least in the context of this thesis where I have persistently stressed the explanatory, as opposed to epistemic, character of Hume science of man, it is persistently neglected in the secondary literature. In keeping with the well established practice of appropriating the support of long-dead philosophical luminaries to one’s position, it was common for logical positivists to interpret Hume as advancing a kind of conceptual scepticism: that is, claiming that distinctively philosophical topics (such as causal necessity) correspond to no aspect of experience and should therefore be dispensed with as cognitively meaningless (see Williams, (2004: 269-270) for overview and Stroud, (1978) for a rebuttal). Similarly, Landy (2009: 9) seems to argue that since Hume doubted the legitimacy of concepts such as necessity, persistence (they cannot be founded upon sense data), Hume denied we possess such concepts at all. This is a non sequitur and is inconsistent with the reading defended here.

from one to another?” (Treatise: 1.3.III, 68). This move amounts to an inversion of his methodological strategy. Instead of searching ‘bottom-up’ for the idea of necessity, Hume is proposing that we approach the matter ‘top-down’: beginning with the transitions we make from what is observed to a causally related unobserved matter of fact in the hope that these will provide a clue as to how the idea is acquired.

It is easy to see, according to Hume, that these transitions are not matters of demonstrative reasoning. The litmus test for such reasoning is whether it is possible to conceive of one object without conceiving of the other. Yet, no object, considered alone, ever supplies a demonstrative reason to infer the existence of any effect distinct from itself, since, as has already been stated, no simple ideas are inseparable. Any cause can be coherently conceived without its typical effect, or with a different effect altogether or none at all. When I see a billiard ball moving towards a motionless one, I can imagine countless scenarios resulting. Although I may conceive that the contact of the two balls will produce motion in the second, I can just as easily conceive both balls coming to absolute rest, or the first ball leaping over the second, or both balls being annihilated altogether.

Far from being a demonstrative matter, it is only by experience that we infer the existence of one object from the observation of another. To be precise, we only pronounce two objects as causally related if we have witnessed instances of their contiguous and successive conjunction in the past: constant conjunction. For example, in the past, a noisy rumbling sound has been constantly conjoined with a large lorry driving past my house, a regularity of experience. One morning I hear such a rumbling sound. Although I am not standing at the window and so cannot see whether there is a lorry driving past my house, I infer its existence anyways as the cause of the sound. As Hume puts it, I “infer the existence of the one from that of the other” and “without farther ceremony call the one cause and the other effect” (Treatise: 1.3.VI, 71). Moreover, the confidence with which we declare this relation corresponds to the wealth of relevant experience available to us. This fact essentially forecloses the possibility we form causal judgements on the basis of our impressions or sense data. This is because experiencing a constant conjunction of resembling impressions does not yield any extra impression over and above the single experience of such a conjunction:

“From the mere repetition of any past impression, even to infinity, there never will arise any new original idea.” (Treatise: 1.3.VI, 72)

Given this, the question is whether it is reasoning which moves us from constant conjunctions to causal judgements, or the imagination. It is here that reason receives its most important hearing. Can it be that people judge two objects to be causally related due to reasoning about their past experience of constantly conjoined objects? If so, what would such reasoning look like? Hume assumes that such reasoning would be based on a ‘uniformity principle’ (UP), stating that future coordination of objects must resemble past coordination of objects. The syllogism would then be:

- (1) In the past, A and B have been constantly conjoined
- (2) The future must resemble the past (UP)
- (3) A
- ∴ B

UP is crucial to the validity of the argument, and so itself needs to be shown to be justified or reasonable. Hume quickly dismisses the idea that UP can be demonstrated, for we can easily conceive of a change in the course of nature and to form a conception is “an undeniable argument” that it is non-demonstrative. The second possibility is that UP is secured on the basis of probable reasoning i.e. inductive reasoning. Such reasoning would presumably be of the form: in the past, the future has resembled the past, and this makes it probable that the future will resemble the past also. However, Hume points out that such reasoning simply presupposes that the course of nature will not change, and so to use it to justify UP would be circular.

Few topics have received more attention than Hume’s critique of inductive reasoning. An influential criticism of this argument is that Hume is wrong to think that reasoning from cause to effect on the basis of past experience of constant conjunctions is possible only if one accepts UP. True, UP is necessary for the *deductive* validity of the inference, but reasoning need not be deductive in order to be warranted.

“...to ask whether it is reasonable to place reliance on inductive procedures is like asking whether it is reasonable to proportion the degree of one’s convictions to the strength of the evidence. Doing this is what ‘being reasonable’ *means* in such a context.” (Strawson, 1952: 257)

In other words, reasoning can be perfectly legitimate without reaching the heights of deductive certainty. It is not my intention here to enter into this debate, but rather to ensure

that this objection does not obscure the central point which Hume, in my view, wishes to extract from his critique of inductive reasoning. This view is explicit in the *Enquiry* when Hume announces that he shall content himself with showing that “our conclusions from that experience [constant conjunctions] are *not* founded on reasoning or any process of the understanding” (*Enquiry*: 4.15, 113). Hume supposes that if our causal judgements were founded upon reasoning, it must be possible to discover the argument thereby employed. The point of his sceptical argument is not that there is no sense in which we might consider causal judgements to be warranted or reasonable, but is rather that the apparent absence of such an argument speaks against the view that our causal judgements are formed on the basis of reasoning:

“[Hume] is arguing that we do not adopt induction on the basis of recognizing an *argument* for its reliability, for the utterly sufficient reason that there *is* no argument... that could have this effect.” (Garrett, 1997: 92)

As such, inductive scepticism is used to support the broader conclusion Hume wants to draw, paving the way for an alternate account of our causal judgements: one which does not trace their genesis to reasoning. Williams (2004: 268) calls this approach ‘methodological scepticism’ since scepticism is utilised as a methodological tool rather than a conclusion in its own right.

This is by no means the only evidence which can be used to support this position. Hume constantly stresses how we seldom reflect upon reasons when we draw causal inferences. Instead, past experience operates on our mind “in such an insensible manner as to be never taken notice of, and may even in some measure be unknown to us” (*Treatise*: 1.3.VIII, 83) and “the custom operates before we have time for reflection” (*Ibid*). Similarly, the fact that our day to day causal inferences are unaffected and unimpaired by sceptical argument, even when those arguments are explicitly endorsed, confirms that these inferences are not delivered by reason. Regardless of whether it is possible to show that our causal inferences are reasonable and legitimate in a non-deductive sense, Hume’s conclusion that the source of our causal judgements is not reason is not significantly threatened.

Instead, Hume decides that we make the transition from observed to unobserved due to the imagination. In particular, the appearance of one object naturally transports us to the idea of its usual attendant because these two ideas have been united in the imagination. The effect of the repetition of resembling instances is a “determination” of the mind to pass from

one object to its usual attendant due to the associative force of the imagination. Consequently, constant conjunction *does* produce something which single experience of conjunction does not, explaining why experience is so important to our causal judgements: this determination of the mind is produced only by the constant conjunction of resembling objects. By “determination” Hume means a feeling of inevitability which accompanies the transition from cause to effect and it is this which is the key to our idea of necessity (Treatise: 1.3.XIV). Hume claims that this feeling is a particular type of impression, an impression of reflection.<sup>33</sup> It is from this impression that the idea of necessity is copied.<sup>34</sup> In this sense, rather than our causal judgements being determined by the prior possession of the idea of necessity, the inverse is correct: the idea arises from the natural propensity or custom of the imagination to move the mind from one object to another.

This still leaves a critical element of our causal judgements unaddressed. Hume takes himself to have shown how, from an idea A, the imagination leads one to the idea B and so on. However, no conditional of this form, even an infinitely long one, can explain why in causal judgements people are led to *beliefs* about unobserved matters of fact.

“...if flame or snow be presented anew to the sense, the mind is carried by custom to expect heat or cold and to *believe*, that such a quality does exist” (Enquiry: 5.8, 123)

Notice that no idea of the imagination can induce a belief in an unobserved matter of fact. While there may be a union in the imagination between my idea of the sun low in the horizon and buildings projecting shadows, the idea of the former cannot alone prompt belief in the latter. It is only an impression or a memory that can produce such a belief. For example, consider the conditional, if the lawn is neat and trim then someone cut the grass ( $A \rightarrow B$ ). If this is to produce the belief about the unobserved matter of fact, that someone cut the grass (B), it is necessary to detach from the conditional, (A). Evidently, only an impression or a memory of the lawn being neat and trim will suffice in this respect. To simply imagine the lawn to be neat and trim will not. This observation holds the key to Hume’s explanation of why we *believe* in the unobserved matter of fact, rather than merely supposing or conceiving it. Recall that according Hume, impressions and memories differ from ideas of the imagination in their force or vivacity.

<sup>33</sup> An impression of reflection is an impression which is caused by something appearing before the mind.

<sup>34</sup> As such, I think that Landy (2009: 9) simply misreads Hume’s position when he says that: “They are the concepts (e.g. necessity, persistence, etc.) that Hume argues we *cannot possibly have* because they do not resemble any of our impressions.”

“...when any impressions becomes present to us, it not only transports the mind to such ideas as are related to it, but likewise communicates to them a share of its force and vivacity” (Treatise: 1.3.VIII, 79)

Where two ideas are united in the imagination, the force of the impression or memory is communicated or transposed to the associated idea. Yet, a belief simply is a lively and forceful idea. In the case of causal judgements then, the force of the impression or memory is transposed onto the associated idea, rendering it a belief in an unobserved matter of fact.

Hume’s position is that our causal judgements derive from a habit or custom of the mind to associate objects which are constantly conjoined. Unfortunately, Hume’s arguments have often been interpreted as aiming to establish a very different conclusion: that there is no physical necessity in the universe:

“Hume is confident that there are and can be no objective physical necessities... in the universe around us.” (Flew, 1986: 5; see also, Kukla, 2011: ii)

Flew notes the “embarrassing consequences” which follow from this, given that many of Hume’s own theories rely upon the notion of causal necessity (*Ibid*). Were this really the conclusion Hume drew from his arguments, the consequences would be very embarrassing indeed. After all, the very idea of a science of man is premised upon explaining qualities of moral subjects by tracing their *causal* genesis. Had Hume really concluded that there are no physical necessities in the universe, he would have had no choice but to condemn his own naturalistic approach as misguided. Fortunately however, this is not Hume’s position: it is not entailed by his conclusion that our causal judgements are the product of custom. Moreover, he explicitly warns against sliding from this psychological conclusion to the metaphysical one Flew critiques:

“I am, indeed, ready to allow, that there may be several qualities, both in material and in immaterial objects, with which we are utterly unacquainted; and if we please to call these power of efficacy, *it will be of little consequence to the world.*” (Treatise: 1.3.XIV, 131)

As Norton (1982: 201) puts it, “[Hume] does not suppose that reality must be of a certain form because a belief that it is so cannot be shaken off, nor does he suppose that there is no reality of a particular sort because the evidence that is available to us falls necessarily short of proving the existence of this particular kind of things”.

## 2.2. *External bodies*

If we turn from the issue of causality to Hume's investigation of another fundamental belief, the existence of external bodies, we find the same basic method and aims present.<sup>35</sup> Hume begins the relevant section of the *Treatise* (I.4.II) with the announcement that:

“The subject, then, of our present enquiry is concerning the *causes* which induce us to believe in the existence of the body.” (Treatise: 1.4.II, 146) [my italics]

That is, the investigation falls within the purview of the science of man. The task is to explain a certain quality of human beings, our belief in external bodies, by tracing its causal genesis. As to the normative credentials of this belief:

“We may well ask, What causes induce us to believe in the existence of body? But it is in vain to ask, Whether there be body or not? That is a point which we must take for granted in all our reasonings.” (Treatise: 1.4.II, 146)

Hume advises that the question of whether our belief in external bodies is correct, whether there *really are* external bodies, is not a topic worthy of consideration given that we have no choice in the matter.<sup>36</sup> Irrespective of whether he is right to think this, it clearly signals the anti-epistemic character of his endeavour.

According to Hume, belief in external bodies should be analysed in terms of two ideas which are “commonly confounded”: the *continual* existence of objects unperceived and the existence of objects *distinct* from their being perceived. These two ideas are to be examined apart. However, such is the intimate connection between them that an answer to why we believe objects to continue to exist unperceived will be an answer to why we believe that they have distinct existence, and vice versa.<sup>37</sup>

As in the case of causation, there are three possible sources of belief to be considered: the senses, reason and the imagination. That the senses cannot give rise to the idea of the

---

<sup>35</sup> By ‘bodies’, Hume means simply those things which comprise the external world. Given this, I shall use the terms ‘external bodies’, ‘external world’ and ‘external objects’ interchangeably.

<sup>36</sup> Evidently, Hume is here pre-empting his conclusion that the belief is a natural belief.

<sup>37</sup> Strictly speaking, this ‘vice versa’ is false. While, from the incontrovertible premise that our perceptions are not continuous, continuity entails distinctness, distinctness does not entail continuity. As Allison (2008: 231) points out, it is conceivable that objects be distinct, and yet be perpetually created. Nevertheless, since the strategy Hume pursues moves from continuity to distinctness, this is not material to the argument.

continual existence of objects is apparent, for this would require the perception of something unperceived. Furthermore, there is little prospect of the senses supplying us with the idea of distinct existence: “a single perception can never produce the idea of a double existence” (Treatise: 1.4.II, 147). Double existence could only derive from some inference of reason or the imagination *beyond* our sensory impressions. Hume reinforces this claim by pointing out that not all of our impressions are ascribed distinct existence. There seem to be three distinct classes in this respect: (a) primary qualities, (b) secondary qualities, and (c) pains and pleasures. Both philosophers and the vulgar (the common man) take primary qualities to have a distinct existence. The vulgar, but not philosophers, take secondary qualities to have distinct existence. And neither the vulgar nor philosophers take pleasures and pains to have distinct existence. Hume is not, at this point, making a claim about the correctness of these classifications, only that (a-c) enjoy parity *as impressions*, and so our beliefs about distinct existence cannot be simply a matter of having sense data.

Building upon this observation, Hume quickly dismisses the possibility of beliefs about external bodies being derived from reason. While philosophers devise arguments supporting their position:

“...’tis obvious these arguments are known but to very few, and that tis not by them that children, peasants and the greatest part of mankind are induced to attribute objects to some impressions and deny them to others” (Treatise: 1.4.II, 150)

By a process of elimination, these meditations leave only the imagination unscathed. Hume’s discussion of how the imagination gives rise to belief in external bodies is exceptionally dense, and in order to keep my discussion as focused as possible, I shall set aside some avenues of inquiry which he broaches only to discard. Following Hume’s own exposition, the skeleton of his argument will be presented first, with the steps explored subsequently.

One of the qualities of the impressions to which external existence is attributed is constancy: resemblance to other impressions over a temporal distance. For example, we find that the perception of the sun or the ocean returns to us after an interval. Even more familiarly, upon closing and opening one’s eyes, we find that impressions prior to closing them resemble those which appear upon opening. On account of their resemblance, we are impelled by the imagination to consider constant impressions *identical*. However, the interval separating the impressions is so contrary to identity that the mind is cast into a state of confused contradiction regarding the matter. In order to resolve this predicament we feign or suppose



that the interrupted perceptions are connected by a continual existence of which we are insensible.

This sketch of Hume's argument is unlikely to look plausible until the key steps are properly explained. The initial or background premise concerns identity, or rather, given that it is temporally spaced (resembling) impressions which are taken to be identical, the argument concerns identity over time. If we did not have such an idea, the imagination could hardly impel us to consider constant impressions identical. However, Hume's early reflections upon this topic are not promising, for he declares that we never perceive an object identical over time. This is because all impressions have an 'internal and perishing existence'; they are bereft of temporal thickness. A single impression therefore can afford only the idea of unity, not identity. On the other hand, perceiving many objects conveys only multiplicity, not identity. Between unity on the one hand and number on the other, there is said to be no medium.

This impasse is resolved by recourse to the imagination. Although strictly speaking, duration entails perceptual diversity and therefore forecloses numerically identity, the imagination is able to engage in a 'fiction' which affords the idea nonetheless: "a single object, placed before us, and surveyed for any time without our discovering in it any interruption or variation, is able to give us a notion of identity" (Treatise: 1.4.II, 155). Hume's thought here seems to be that the idea of identity derives from a foible in our imaginative capacities. In particular, we have the idea of identity because we can conceive of, or 'keep in view', an object over time without discovering or noticing that it varies, despite its *actual* variance. So for series of resembling impressions, symbolised as:

AAAAAAAAAAAAA,

The resemblance and uninterrupted nature of the sequence is sufficient for the imagination to prompt the idea of identity over time. Identity is therefore defined as "the invariableness and uninterruptedness of any object, thro' a supposed variation in time" (Treatise: 1.4.II, 156).

Having secured the idea of identity, Hume needs to show why the imagination forces us to ascribe identity to constant impressions, which are invariable (a strong form of resemblance), but interrupted. The reason for this, Hume claims, is that the disposition or act of surveying a sequence of constant impressions 'feels' like the perception of an identical object. The passage from one resembling idea to another is so smooth that we confound the

constancy of our ideas for their identity: the invariable and uninterrupted view of an object. In other words, a series of constant impressions:

(1) AAAABBBBAAAAA

Places the mind in almost exactly the same dispositional state as an uninterrupted series of invariable impressions:

(2) AAAAAAAAAAAAAA

Since the imagination compels us to regard cases of (2) as cases of identity, by confounding (1) with (2) so too we are inclined to regard constant impressions as identical.

*However*, despite this tendency to inflate constancy into identity, the interruption of constant impressions is contrary to the idea of identity, suggesting instead diversity. As such, we find ourselves caught in a *dilemma or contradiction*: impelled to ascribe identity to constant impressions and yet unable to do so. This state of uneasiness and perplexity “produces a propension to unite these broken appearances by the fiction of a continued existence” (Treatise: 1.4.II, 159). That is, we overcome the contradiction by supposing that objects have a continual existence beyond their being perceived, and so secure their identity over time in a way which is compatible with their perceptual intermittency.

Having uncovered the emergence of the idea of continued existence, Hume quickly explains why we come to believe it, how it gains the requisite FLV. Although up to this point speaking loosely of the ‘recognition’ of resembling impressions or ‘holding an object in view’, it is clear that this capacity depends upon the memory. For example, a sequence of constant perceptions,

AAAABBBBAAAAA

can only be recognised as constant (invariable, but interrupted), insofar as we can recall that the perceptions prior to the interruption resemble those that follow it. As already noted, ideas of the memory are the functional equivalents of impressions according to Hume, bearing comparable amounts of FLV (in contrast to ideas of the imagination). Hume concludes from this that since our propensity to feign the continued existence of objects arises from “lively

impressions of the memory”,<sup>38</sup> these impressions must bestow vivacity upon the fiction, rendering the fiction a belief (Treatise: 1.4.II, 161).

There are two serious objections to Hume’s explanation of our belief in external bodies. The first concerns how the idea of continued existence is acquired. The problem is that, upon closer inspection, Hume does not have an account of this acquisition at all, let alone a plausible one. To illustrate this, it is worth reviewing the relevant step in Hume’s argument, the point at which the idea of continued existence is supposed to emerge:

“The smooth passage of the imagination along the ideas of the resembling perceptions makes us ascribe to them a perfect identity. The interrupted manner of their appearance makes us consider them as so many resembling, but still distinct beings, which appear after certain intervals. The perplexity arising from this contradiction produces a propension to unite these broken appearances by the fiction of a continued existence, which is the third part of that hypothesis I proposed to explain.” (Treatise: 1.4.II, 159)

There is a state of perplexity or contradiction, and this prompts the idea of continued existence which resolves the contradiction. In other words, the acquisition of the idea is explained by the instability or perplexity of the mind prior to its acquisition. We acquire the idea *because we have need of it*. Yet, this is surely one of the weakest claims to be found in the whole of the *Treatise*. As Stroud puts it witheringly:

“Acquiring a new idea cannot be explained as simply a matter of selecting, among a number of antecedently intelligible alternatives, the one that best squares with all the available data, since in acquiring a new idea we come to find something intelligible or to understand something that we did not understand before.” (Stroud, 1977: 108)

Indeed, a much more coherent reconstruction of this stage of Hume’s argument is that it is not directly concerned with the idea of continued existence at all, but with why we come to believe or judge an idea already apprehended. For one thing, Hume’s own language suggests this. The state of perplexity is said to prompt “a propension to unite these broken appearances

---

<sup>38</sup> Although strictly speaking the memory is a faculty of *ideas* not *impressions*, Hume talks of ‘impressions of the memory’ in places, apparently to reinforce the functional equivalence of ideas of the memory and impressions.

by the fiction”, “our inclination to that supposition” or the “feigning a continued being” (Treatise: 1.4.II). This sounds much more like the formation of a belief, albeit one which perhaps involves a certain degree of self-deception, than the acquisition of an idea. Moreover, tending towards belief formation makes sense given that only the belief in the continued existence of objects, as opposed to the mere idea, would be capable of resolving the state of perplexity described.

For example, imagine that I have spent years working on a paper. This paper is not merely an exposition of a view I happen to regard as correct or true, but one which captures the essence of all my philosophical thinking for years previous. However, the paper is rejected by every journal that I send it to outright. What am I to think of the paper? On the one hand, to discard it would be to discard the conviction that I am good at what I do, that my career is being spent in a productive fashion, and so all of my instincts of self-worth and self-preservation impel me to regard it as a piece of great insight and importance. On the other hand, it has been rejected by my peers. This casts me into a state of contradiction: inclined and yet unable to view my paper as philosophically revelatory. In order to escape this perplexing state, I feign or suppose that the journal reviewers have an unreasonable prejudice against the position defended in the paper. This feigning reconciles the need to regard the paper as brilliant with the rejection it has suffered.

In this example, it is clear that the state of perplexity or contradiction is resolved by a belief. When I feign that the reviewers are prejudiced, I must really affirm or assent to it if the perplexity is to be dispelled. Merely having the *idea* of prejudiced reviewers could not serve in this respect. Similarly in the case of the continued existence of bodies, it is the belief that bodies continue to exist which is capable of resolving the contradiction. In other words, at this stage of his argument, Hume is forging a connection between a natural state of perplexity and a belief in the continued existence of bodies: the latter resolves the former and the former explains the latter. However, this serves to underline the objection: this is not an account of the acquisition of the idea of continued existence at all. By Hume’s own lights, all belief presupposes the idea believed.<sup>39</sup> In the paper case, I cannot form the belief that the reviewers are prejudiced if I lack the idea of prejudice. Equally, it is no good being told that we believe bodies to have continual existence because this resolves a contradiction, since this

---

<sup>39</sup> Presumably this is why Hume never uses the terms ‘belief’ and ‘judgement’, but masks the true character of this stage of his argument with talk of ‘feigning’ and being ‘inclined to suppose’.

presupposes the idea of the continual existence of bodies and this is precisely what is supposed to be under investigation!

The second objection to Hume's account is that he is unable to give an adequate description of the original position from which the belief in external bodies originates.

“Whoever wou'd explain the origin of the common opinion concerning the continu'd and distinct existence of body, must take the mind in its common situation.” (*Treatise*: 1.4.II, 164).

It is certainly correct that tracing the genesis of the judgement presupposes an account of the common situation from which it springs. Evidently, the minimal standards of adequacy for such an account is that it be free from the judgement in question. It is of no explanatory value to trace the belief in external bodies back to a common situation which presupposes, in some respect, the relevant judgement. In the case of causation for example, Hume had such a coherent description of the common position, one involving the constant conjunction of impressions and laws of association. In the case of external bodies however, the matter is not so straightforward.

Hume labels the common situation the ‘vulgar position’, by which he means to contrast it with the position held by philosophers. The crucial element of the philosophical position is that it invokes a ‘double existence’ of perceptions and objects. On the one hand, we are only ever directly acquainted with our own perceptions. Yet equally, on this view, there also exists a realm of objects which our perceptions resemble. Hence, philosophers recognise a dichotomy of fleeting perceptions on the one hand and objects which are coexistent and resembling on the other.

By contrast, Hume insists that the vulgar embrace no such view. They “never think of a double existence internal and external, representing and represented” (*Treatise*: 1.4.II, 158). Apart from the plain conviction that this is just ‘not how it is’ with the common man, Hume offers two reasons in support of this. Firstly, if the common situation involved a dichotomy of perceptions and objects, the idea of the external existence of bodies would already be present since this simply *is* a dichotomy of perceptions on the one hand, and objects which are distinct from those perceptions on the other. Secondly, Hume argues that were the common situation the philosophical one, belief in the existence of bodies would never emerge. This is because we would never be tempted to ascribe identity to our perceptions if we recognised

them *as* perceptions – internal, fleeting images. In other words, the philosophical position is subsequent to belief in the existence of bodies:

*Common situation/Vulgar position* → *Belief in external bodies* → *Philosophical position*

Instead of drawing a philosophical distinction between perceptions and objects, the vulgar understand that, “those very sensations, which enter by the eye or the ear, are with them the true objects, nor can they readily conceive that this pen or paper, which is immediately perceived, represents another, which is different from, but resembling it” (Treatise: 1.4.II, 156). But what does this amount to? The most natural way to understand this is that the vulgar are direct realists about perception: they take themselves to be directly acquainted with pens, paper and teapots. But this surely entails that they already have the idea of, at least, the distinct existence of bodies, if not that of continued existence also, because that is what pens, paper and teapots are: things which exist distinctly! A clear way of presenting this difficulty is to apply it to a particular step in Hume’s account, one explanatorily prior to the emergence of the idea of continued existence i.e. prior to the state of contradiction and perplexity.

Suppose that I am standing atop a castle turret and have before me a glorious vista across open countryside. I close my eyes and when I reopen them, the same vista reappears to me. Hume says that in this case, there is a natural relation of resemblance holding of my impressions in the two instances. Because a sequence of constant impressions ‘feels’ the same as the uninterrupted view of an invariable object, the imagination impels me to construe them as identical. Now, this is the pivotal question: *what is it* that I am inclined or impelled to consider identical? The obvious response is that it is the countryside, with its rolling hills and scattered trees which is taken to be identical or the same. Yet, this cannot be right for it would amount to saying that what is the same is an object or collection of objects distinct from my perception and this idea of distinctness is not available at this point in the account.

Since it cannot be distinct objects which are considered identical, it is tempting to ascribe to the vulgar something of the philosophical view Hume denies they have. According to this, rather than it being hills and trees which are deemed identical, it is only our *perceptions* which are deemed identical. Indeed, Hume suggests as much when he says that:

“The very image, which is present to the sense, is with us the real body; and ‘tis to these interrupted images we ascribe a perfect identity.” (Treatise: 1.4.II, 158)

However, this is equally implausible given Hume's belief that if the common situation of the vulgar made a distinction between our perceptions and objects, the idea of the continued existence of bodies would never arise. Indeed, it seems more appropriate to interpret Hume's loose talk of 'ascribing perfect identity to (perceptual) images' as a tacit recognition of the conundrum he faces: namely, that 'ascribing perfect identity to objects' is equally problematic.

### ***2.3. The thread from Hume's science of man to the Critique of Pure Reason***

Hume's failure to explain the genesis of our belief in external bodies is extremely significant, much more so than it first appears. Hume presents the issue of external bodies as one among the many which a science of man must address: other similarly fundamental judgements such as those concerning causation, personal identity etc., also demand attention. Viewing these as relatively distinct topics for investigation suggests that difficulties with one of them need not afflict the others. Roughly, just because Hume fails to provide an adequate account of the belief in external bodies, this need not permeate the broader naturalistic endeavour.

However, this rather benign assessment is over-simplistic. The problem is that belief in external bodies is not isolated in the way that this analysis would suggest. What is in question is not a single, confined belief, 'Distinct and continuous bodies exist', nor even a broad class of beliefs of the form 'X exists distinctly and continuously'. In fact, belief in external bodies is implicit in *all* judgements. As I pointed out earlier, 'external bodies' for Hume is interchangeable with 'external objects' or even the more ontologically neutral 'external world'. By it, he simply means to denote whatever is understood to exist distinct of our perceiving it and exist continuously when unperceived, and this can be construed broadly to include objects, properties, relations etc. (as far as I can see, it can include whatever ontological categories one may endorse). Yet, it follows from this that most, if not all of our judgements invoke or presuppose belief in external bodies, for this is to say no more than that our judgements are directed at the world: that they have accuracy or veridicality conditions which are satisfied, or not, by the way the world is, rather than by the way our perceptions are. For example, when I judge 'this rose is red', I am not claiming that there is a connection holding of my ideas/perceptions. Rather, I am attributing a property of objects which exist distinctly and continuously to an object which exists distinctly and continuously. The same essentially applies to: 'swans are white', 'the motion of the billiard ball caused the second

billiard ball to move', and 'I am identical with myself'. In other words, to ask why we make judgements about external bodies/object/existence is to ask why judgements have *intentional content*.

That Hume did not fully appreciate this is manifest in how he attempts to explicate the common position from which judgements of external bodies emerge. In particular, he treats this as a state in which many judgements – suitably forceful ideas - are possible, just not that of external existence. Most notably, he declares that subjects in the common position make judgements about identity. Yet, this is to neglect how *all* judgements seem to be judgements of external bodies. Hume's neglect of this point made it all but inevitable that his account of the common position from which belief in external bodies springs would presuppose what is in need of explanation.

Since belief in external bodies is presupposed by all judgements, what appears to be an isolated shortcoming of Hume's science actually comprehensively pervades it. This is because, recall, the explananda for a science of man are the judgements (or at least, an important subset thereof) people form. The whole approach is founded upon treating judgements as a data set peculiar to moral subjects, to be explained by tracing their genesis back to more primitive causal antecedents. Consequently, in failing to account for the belief in external bodies, Hume fails to account for the most fundamental feature of this data set – that they are intentional. He does not explain why judgements ubiquitously have intentional content.

In the context of Hume's theory of ideas, this worry could be posed more generally as a challenge for this framework. Insofar as this is the framework within which the judgements of moral subjects are to be explained, a scientist of man is faced with a pressing challenge:

*The Intentionality Problem:* How is it that, by means of our perceptions, we enjoy mental states which so much as *seem* to be about or directed at things other than these perceptions themselves, specifically objects which exist distinctly and continuously?

Given that this question is simply a theory of ideas-relative formulation of the question 'how is it that judgements are objective / why do humans form belief in external bodies?' Hume, inevitably, has no adequate response to offer here.<sup>40</sup>

---

<sup>40</sup> Consider how, as I have pointed out on numerous occasions, impressions and ideas have content essentially according to Hume. They are *inherently* in the representation business. But to say this much is already to



It is important to emphasise that presenting this as a problem for Hume does not conflict with my earlier pronouncements in Section 1.4 and 1.5 regarding the naturalistic status of his theory of ideas. Recall, I argued there that Hume's theory of ideas is primarily intended as a psychological thesis, an inference to the best explanation, rather than an epistemological thesis. I argued that although the theory of ideas raises epistemological questions, namely whether our mental states successfully make cognitive contact with the world, these are questions Hume legitimately considers to fall outside the purview of a science of man. Now, one possible suggestion at this point, and one which I am anxious to guard against, is that the intentionality problem is essentially an epistemological challenge of the sort I earlier declared Hume need not concern himself with. Were this the case, I would clearly be contradicting myself, but it isn't, so I'm not. The Intentionality problem is not that of determining whether our representations or states are veridical or accurate. Nor is it that of determining criteria by which we might justifiably take them as veridical. Rather, it is the challenge of determining why it is that they are directed upon or are about a world which exists distinctly and independently at all: a problem of intentionality. Why they so much as *purport* to be about the world in this fashion (even if they are actually *non-veridical*).<sup>41</sup> Insofar as this is simply a theory of ideas-relative formulation of Hume's question: 'how do we acquire belief in external bodies', it certainly is a problem for a scientist of man. As we might say, the intentionality problem is not that of whether our states are veridical or not – a problem which falls outside the purview of Hume's approach –, but rather, the problem of why our states so much as purport to veridicality, a problem which absolutely does fall within the purview of Hume's approach.

In my view, this intentionality problem marks the most fundamental connection between Hume's science of man and Kant's transcendental philosophy. However, whereas Hume was never properly cognizant of the issue, Kant has a sophisticated grasp of it, placing it at the heart of the first *Critique*:

“How does it come about that we posit an object for these representations, or attach to them, beyond their subjective reality as modifications, some kind of an objective reality?” (A197/B242)

---

attribute to subjects the idea of external bodies, and at least some beliefs in external bodies. There is a certain sense then in which Hume's pursuit of the question, 'why do moral subjects form beliefs in external bodies' is confused by his own standards.

<sup>41</sup> I discuss this distinction in more detail in relation to Kant in Section 4.2.

In this sense, Kant picks up a challenge which is of pressing concern for a science of man - albeit one which Hume overlooks - and attempts to answer it from within his own distinction philosophical standpoint.<sup>42</sup> The account Kant develops will be the primary concern of Part II, but I want to conclude my discussion of Hume by foreshadowing why this account is so important, and why it is an appropriate topic to turn to having explored Hume's science of man. The significance of Kant's account lies *not merely* in that he is grappling with a problem which should be on Hume's horizon also, although this certainly invests it with enough significance to be getting on with. In the context of this thesis, the reason why Kant's response to this challenge is so consequential is that fundamental to the account he develops is the role he assigns to *judgement*. Indeed, Kant's conception of judgement only makes sense in light of his interest the problem of intentionality; they are inextricably entwined for him. This is of pivotal importance: it ensures that a philosophical thread which moves from Hume-on-judgement (or more generally, judgement within a science of man - see below) to Kant-on-judgement is much more than a survey of disparate and unconnected approaches to judgement: much more than an inspection of a fraction of Sluga's "hyper complex" domain. The problem of intentionality binds together these approaches relative to a common philosophical question, providing the thread or philosophical narrative which runs throughout. While the theme of this thesis is *judgement*, the integrity of this theme derives from its being interwoven with the problem of intentionality.

---

<sup>42</sup> Longuenesse (1998: 186-7) considers, and then rejects, interpreting Kant's refutation of Hume in this way. The reason for this is that "of course, Hume does have a story to tell about how we come to believe that our impressions and ideas represent properties of objects "whatever the state of the subject"". However, as I have argued, not only does Hume lack an adequate story, he has only a weak grasp of the problem.

## Chapter 3. Cognitive Science

---

My concern in this chapter is to show how Hume's naturalism forms the methodological framework for contemporary cognitive science and how, as such, cognitive scientists can be seen reinvigorating the science of man, two hundred and fifty years after the publication of the *Treatise*. I shall begin by observing some general points of comparison, before turning to two case studies of how cognitive scientists have approached grammaticality judgements and moral judgements.<sup>43</sup>

### 3.1. Picking up humean strands

Hume's attempt to chart a "[human] mental geography, or delineation of the distinct powers of the mind" (Enquiry: 1.13, 93) has been widely identified as a forerunner of contemporary cognitive science (e.g. Chomsky, 2000: 164; Hinzen, 2006: 1). Indeed, Fodor (2003: 134) claims that the *Treatise* is the "foundational document" of cognitive science. Unfortunately however, this connection is too often amalgamated into, or subsumed under, a more general historical reading, according to which the whole of early modern philosophy is portrayed as a fledgling empirical science of the mind. Hatfield notes how it is increasingly common to see the early moderns "as having been pioneers in the development of a naturalistic approach to the mind, and perhaps to have been early labourers in the field of cognitive science" (Hatfield, 1990: 10).<sup>44</sup> On this view, early modern philosophers were primarily interested in what would today be thought of as largely psychological problems: how does the mind work? What are its component parts? For example, Chomsky has repeatedly insisted that the common practice of labelling the deep shifts in intellectual thought which laid the foundations for cognitive science in the 1950s the 'cognitive revolution' is misleading. Instead, he suggests, 'the second cognitive revolution' would be more appropriate since cognitive scientists have been "rediscovering" themes which first arose towards the end of the seventeenth century

---

<sup>43</sup> Since my case study of moral judgements is more original and less familiar than that of grammaticality judgements, I shall dwell on it longer.

<sup>44</sup> I do not mean to suggest that this is the overriding view of the early modern period amongst, say, historians of philosophy or, indeed, philosophers at large. However, where the connection between Hume and cognitive science has been identified, often this has been within the confines of this broader historical stance.

(Chomsky, 2007: 3). As such, to the extent that philosophy nowadays eschews psychological questions, it is seen to be betraying its early modern roots (e.g. Knobe, 2007).

Essential to this exegetical strategy has been to interpret the early modern preoccupation with the existence of innate ideas, a topic which enlivened and animated the likes of Locke and Leibniz, as a precursor to current debates in the field of developmental psychology. These debates centre upon the question, ‘to what extent, and in what sense, are our mental powers genetically constrained?’ with a scale of answers possible. At one end of the scale, a strong empiricist contends that our genetic endowment greatly under specifies the type of mental faculties which humans acquire, providing little more than a general learning device, enabling abstraction and generalisation. At the other end of the scale, a strong nativist favours a substantial genetic endowment which severely restricts the space of possible (that is, biologically possible) psychological development. Such developmental questions are matters for empirical science if anything is and by equating these questions with the early modern concern with innate ideas, the broader historical outlook is largely secured. For example, Steven Pinker (2002: 18-19) interprets Locke’s famous proposal that the mind is a *tabula rasa* as an early example of psychological behaviourism. Specifically, Locke is said to be denying that is such a thing as human nature: that there are virtually no initial constraints upon the pathway of mental development. Psychological behaviourists in the twentieth century simply updated this view by “replacing” Lockean ideas with stimulus and response, and association with conditioning.

However, portraying Locke as an early psychological behaviourist is plainly inconsistent with much of what he says in the *Essay Concerning Human Understanding*. For example he says that “The candle, that is set up for our purposes, shines bright enough for all our purposes”, adding that objects must be entertained in that way that is suited to our faculties (Locke, 1690: 13). Locke accepted that there was such a thing as human nature, constantly stressing that we are not angels and can only make best use of the limited faculties we have available to us. As Blackburn (Ms.) states in his review of Pinker, “Locke only wanted to deny innate ideas and knowledge, not innate powers or tendencies, nor innate limitations, nor innate cognitive and emotional capacities”. For Locke, the *tabula rasa* metaphor was not a rejection of a psychological thesis regarding mental development, but a rejection of an epistemological thesis. Specifically, Locke was claiming that what can be

known prior to experience is highly limited.<sup>45</sup> He thought that aside from a class of intuitive and demonstrative propositions, there are no propositions we are justified in asserting irrespective of sense data. Whatever psychological nativism one may choose to endorse, there is no nontrivial move from this to any claim about what we are *warranted* in asserting a priori;<sup>46</sup> it is the latter issue which Locke is interested in: “The initial and abiding focus of Locke’s investigation is the ground of belief... rather than the naturalistic search for causal mechanisms... or psychogenesis” (Hatfield, 1990: 29-30).<sup>47</sup> Indeed, it is precisely this epistemological question which Hatfield (1990; 1997) persuasively argues was the centrepiece of early modern philosophy, one which concerned the *real use of the intellect*. Many philosophers believed that, alongside a logical use whereby ideas were compared and contrasted, the intellect also had a ‘real use’ in which it operated as an *epistemic power* (Hatfield, 1990: 3). According to this view, the intellect, properly employed, was capable of delivering substantial truths about the world a priori, without the aid of the senses. It was this real use of the intellect which Locke sought to curtail with his *tabula rasa* metaphor. By contrast, for Descartes the undiluted operation of the intellect, ‘the clear and distinct light of reason’, was capable of yielding considerable knowledge of the world e.g. the essences of substances. Cudworth believed that the real use of the intellect had a very wide range. Most famously, he argued (1731: IV.2, 83) that our moral knowledge, that of “virtue, vice, honesty, dishonest, justice, injustice”, was derived not from the senses, but secured by the activity of the intellect. Crucially, the key claim for Cudworth was not merely that our moral ideas are innate, but that they provide us with secure *knowledge* of an independent moral reality i.e. knowledge of matter of fact.<sup>48</sup> As Norton (1982: 31) comments, “There can be no question, either, of this moral reality being [for Cudworth] mutable, private, or conventional. Knowledge is, after all, of that which is unchangeable and public, and of the nature of

---

<sup>45</sup> For Locke, what can be known prior to experience is either a matter of intuition or demonstration. Intuitive knowledge involves comparing our ideas and recognising relations they bear to each other e.g. white does not resemble black. Yet, such knowledge does not extend our knowledge of the world, only of our ideas (it is, roughly, analytic in Kant’s terminology). Demonstrative knowledge is a matter of moving from one idea to another by interposing a chain of related intermediaries. According to Locke, this includes mathematics, the existence of God and key moral axioms (cf. Allison, 2008: 65-72), - knowledge sufficiently extensive to call into question the appropriateness of calling Locke an ‘empiricist’ at all.

<sup>46</sup> This is for the obvious reason that the ‘innate’ basis of a belief does not preclude error. For example Richard Joyce (2007) argues (a) that our moral concepts are innate (or evolved) and, (b) that we are not justified in endorsing any moral proposition. There is certainly nothing incoherent about such a view.

<sup>47</sup> By contrast, “psychogenesis” is *precisely* what Hume is interested in.

<sup>48</sup> For this reason, hypotheses regarding the existence of a genetically encoded language acquisition device would have been of little interest to Cudworth (pace Chomsky, 1966: 100).

things.”<sup>49</sup> While the early moderns were keenly occupied with the matter of innate ideas then, this was understood within this broader epistemological frame, foreclosing any simplistic ‘psychologistic’ readings of the period.

It follows from this that it is imperative to hold apart the claim that Hume’s science of man is a forerunner to contemporary cognitive science from the claim that the early modern period as a whole should be read as a kind of nascent psychology. For while the latter claim is mistaken, in Hume’s science of man, something, recall, he considered to be “almost entirely new” (Treatise: intro., xvi), we *really do*, I think, find the seeds of cognitive science. This marks Hume’s project out as unique, and renders more valuable an assessment of its relation to cognitive science. Although this is not the place for a comprehensive discussion of the history and philosophy of cognitive science, I think that some brief observations regarding its foundations can serve to illuminate this relation. Consider the following:

“It is the business of normal empirical science – of the burgeoning field of “cognitive science,”... to generate best hypothesis about what states and processes would best explain the creatures’ observable behaviour.” (Biro, 1993: 52)

In this quote, Biro intentionally describes cognitive science in terms which can be accurately applied *verbatim* to Hume’s naturalistic approach in order to underline the parallels (these parallels are the subject of the article in question). Yet, in fact, the language should be stronger. Specifically, while Hume talked of targeting ‘the observable behaviour’ of moral subjects, but largely concerned himself with certain universal judgements produced regarding matters of fact, so too linguists and moral psychologists have had a special interest in explaining the *judgements* people make. For example, Baker (2008: 235) stresses that essential to the foundation and subsequent flourishing of linguistics in the 1950s was the identification of a relatively narrow and well-defined linguistic explananda: subjects’ grammaticality judgements. The challenge was (and still is) to account for how native speakers judge certain strings to be linguistically acceptable or sentences of the language. Crucially, isolating acceptability judgements in this way entails the eschewal of other possible, and perfectly coherent, linguistic explananda, most notably, linguistic performance or language use: why speakers say what they say in certain circumstances, rather than

---

<sup>49</sup> On Norton’s reading, the Scottish Naturalists, Kames, Reid and Beattie, were also appealing to this real use of the intellect, i.e. an *a priori* warrant for non-demonstrative truths, when they spoke of ‘natural belief’ (cf. 1982: Chapters. 4 & 5). This reading strikes me as plausible and shows that their criticisms of Hume were largely misplaced.

something else (or indeed, nothing else) (cf. Newmeyer, 1986: 3). By contrast, accounting for linguistic performance had been the major desideratum for earlier behaviourist sciences of language. For example, Skinner (1957) had aimed to explain (and so predict) the linguistic behaviour or performance of speakers, that is, their utterances, on the basis of environmental stimuli and their history of conditioning. Chomsky (1959: 55) influentially argued *not* that such a desideratum is incoherent or ‘not really linguistic’, but simply that it is too difficult to be the subject of scientific investigation, a state of affairs which has yet to change (Chomsky, 2000: 17; Mukherji, 2010).

Just as in linguistics acceptability judgements are the foundational explananda for the discipline, in moral psychology subjects’ moral judgements have played this role. That is, peoples’ moral judgements constitute, for the moral psychologist, the primary data set: “a set of considered judgements” which are stated “as explanandum sentences” (Mikhail, 2009: 27-28). It is this data for which a principled account is sought. By contrast, moral psychologists have largely forborne from investigating other moral phenomena, such as moral *behaviour*. Just as linguists consider linguistic performance to be too complex a matter for the current state of inquiry, so too in the corresponding case of moral behaviour (Dwyer, et al., 2010: 489). As Abend (2012: 1-3) puts it, in moral psychology, ‘judgement-centric approaches’ dominate.

Having isolated a class of such judgements, understood as an identifiable kind of observable behaviour, the task for the cognitive scientist is to trace their psychological *genesis* or “causal etiology” (Dwyer, 2009: 279). That is, judgements are treated according to the ‘science of man template’. Judgements are the *explananda* of inquiry. In my view, this point of convergence is sufficiently critical to justify the claim that linguists and moral psychologists are pursuing a science of man.

Moreover, the parallels between Hume’s naturalistic approach and contemporary cognitive science go further. For Hume, tracing the etiology of judgements meant tracing their causal ancestry back to more primitive mental elements: impressions, ideas and natural relations holding thereof. This theory of ideas was the theoretical framework within which the explanatory task was addressed. In turn, it is possible to ask what the corresponding theoretical framework within which cognitive scientists address the question of causal etiology is. According to Bermudez (2010: 6), “the guiding idea of cognitive science” is that mental activity is information processing. It follows from this that insofar as the judgements

people draw, grammatical, moral or otherwise, are to be amenable to a cogno-scientific investigation, that is, insofar as they are suitable explananda, they must be produced by mechanisms or systems which process information. Yet, ‘information processing’ is vague if anything is, and a more concrete formulation of the notion of information processing is that the relevant mechanisms or systems are computational.<sup>50</sup> This computational paradigm has been central to developments in both linguistics and moral psychology:

“The general idea is that the language faculty involves a precisely articulated computational system.” (Chomsky, 1986: 204)

“EMP [(empirical moral psychology) targets the computational mechanisms and cognitive architecture that is responsible for the explicit intuitions [judgements] about justice, rights and welfare that arise as people attempt to resolve moral quandaries.” (Huebner, 2011: 52)

Evidently, the full breadth of this topic extends well beyond what can be broached here. Nevertheless, saying a bit more will indicate why this computational paradigm is relevant to a comparison of Hume’s naturalism and cognitive science. Marr (1982) influentially argued that any computational system can be described at three levels of analysis:

1. Functional level: Describes the function or processing task performed by specifying the input to the system and the output of the system.<sup>51</sup>
2. Algorithmic level: Describes how the system performs the function described in (1) by stating the algorithms according to which the input is mapped to the output.
3. Implementation level: Describes how the execution of the algorithms described in (2) is realised by or implemented in a supervenience base e.g. neural circuitry (cf. Bermudez, 2010: 47-54 & 127-134).

According to Marr, accounts ought to begin at the top level of description, i.e. (1), since the higher order descriptions constrain the descriptions given at the level(s) below. Perhaps most

<sup>50</sup> Strictly speaking, information processing is a broader notion than that of computation. Although Fodor once claimed that the computational paradigm is the ‘only game in town’, connectionism is generally seen nowadays as a viable alternative (though, for an influential denial of this point, see Fodor and Pylyshyn, 1988). Either way, since both linguistics and moral psychology have been founded upon the computational paradigm, I shall set aside connectionism here.

<sup>51</sup> Arguably Hume grasped the importance of this functional level of description. Compare it with Stroud’s (2006: 350) assertion that Hume’s enterprise is “illuminating in direct proportion to the size of the gap it posits between the impoverished non-propositional *input* human beings receive and the profuse and highly articulated *output* of beliefs and other attitudes they end up with” [my italics].



importantly, it is possible to describe a cognitive system at the Functional and Algorithmic levels, without prejudicing, or independently of, the question of how the input, output and algorithms of the system are implemented or realised i.e. descriptions at level (3):

“It is possible to study formal algorithmic processes without regard to how such processes are *physically* instantiated in an actual device.” (Pylyshyn 1980: 115)<sup>52</sup>

As such, it is possible to study computational systems in abstraction from their physical implementation: namely, at the Functional and Algorithmic levels of description. The notion of an algorithm is that of a purely mechanical procedure: specifically, a finite set of rules that are unambiguous and can be applied to objects and yield objects as an output. That is, the Algorithmic level describes functions which map the inputs of a system, as per described at the Functional level, to the outputs of the system, as per described at the Functional level. This mapping is a kind of process, and insofar as the relevant inputs and outputs are appropriately understood as information, this computational paradigm is an information processing paradigm. In this way, the inputs and outputs of the system, the bearers of information, are the primitives of the system, the units over which the algorithms are defined. In cognitive science, these units are taken to be mental symbols or representations. Insofar as they bear information, they have certain semantic properties, and insofar as they are algorithmic primitives they have also certain formal or syntactic properties.

To a certain degree, the details of this picture do not matter too much. What does matter is that insofar as cognitive scientists have sought to account for their explananda by recourse to a computational paradigm, they necessarily invoke a domain of mental representations. These mental representations are postulated as part of an inference to the best explanation, and feature in an account of the causal genesis of certain judgements. The parallels to Hume’s own theory of ideas should be obvious. For instance, compare Hinzen’s (2006: 42) assertion that

“Chomsky’s study of the mind... is crucially not concerned with mental *representation*-as a relation between mind and world...-as with mental *representations*, considered as specific natural objects that enter into language use and are invoked to explain it.”

---

<sup>52</sup> This foundational claim is one which philosophers of mind have not always been sufficiently sensitive to, occasionally resulting in misplaced criticism of certain cognitive scientific proposals e.g. Johnson, (2012). See Kirkby (2013) for a critique of Johnson on this point.

, with:

“...he [Hume] treats all cognitive operations... as consisting entirely of operations performed with and on mental representations” (Garrett, 1997: 39)

Admittedly, this parallel should not be overstated. Insofar as Hume had no concept of computation or of an algorithm, he understands mental representations very differently. For Hume, mental atoms or representations are understood in imagistic terms. Perceptions are simply mental images of some degree of clarity and vividness. It never occurs to Hume that over and above their imagistic properties, perceptions may also have formally specifiable, syntactic properties. Of course, insofar as the idea of syntactic properties only really arises with that of computation, this is hardly surprising. Nevertheless, since construing mental representations as syntactic objects has been critical to cogno-scientific accounts, this still marks a significant disanalogy between the two conceptions of mental representations.

Furthermore, while both Hume and cognitive scientists appeal to mental atoms as part of an inference to the best explanation, there is a radical divergence in the range of relations these atoms are subject to. Hume avails himself of remarkably few resources in accounting for the judgements of moral subjects. Three natural relations, resemblance, contiguity and cause and effect hold of ideas and serve to guide the primitive act of the imagination: association. Hume’s explanations for judgements of any kind appeal to little more than the ways in which association, guided by these qualities, conflates our ideas and is generally overactive like ‘a ship set in motion’. As such, the relations mental representations are subject to are highly circumscribed and completely domain general. By contrast, the modularity of the mind has long been an influential doctrine in cognitive science: where the mind is comprised of a number of domain specific systems (Fodor, 1983). An effect of modularity is that the relations which mental representations are potentially subject to are far more numerous than those considered sufficient by Hume.

Despite these significant qualifications, the points of comparison between Hume’s naturalistic approach and cognitive science – the common conception of judgement and the postulation of a realm of mental representations as an inference to the best explanation – warrant the assertion that cognitive scientists are engaged in a science of man. I now turn to case studies of contemporary linguistics and moral cognition.

### 3.2. *Grammaticality judgements*

The point of departure for linguistic inquiry is the judgements speakers of natural languages make about physical strings. For example:

1. John likes him.
2. John thinks that Bill likes him.
3. Every man loves some woman.
4. Loves some woman every man.

Speakers of English have clear judgements about these strings. In (1) ‘him’ cannot referentially depend upon ‘John’. In (2), ‘him’ can referentially depend upon ‘John’, but need not. (3) is ambiguous; it can mean either that there is a particular woman that every man loves, or that every man loves a woman, but not necessarily the same one. (4) is unacceptable. In other words, these are judgements about how sounds (or gestures or marks) can and cannot be paired with meaning. Certain connections of sound and meaning are acceptable, others are not.<sup>53</sup> Such corpuses of acceptability or grammaticality judgements are the explananda for linguists, that which grounds and motivates the inquiry. Equivalently, acceptability judgements are understood as a kind of observable behaviour for which principled explanation is sought. This principled explanation is a matter of tracing the causal genesis of the judgements in question.

What kind of competence must be attributed to speakers in order to account for, or explain the grammaticality judgements they produce?<sup>54</sup> It should be clear that invoking folk psychology is hopeless in this regard. Notoriously, people have sparingly few beliefs about grammar, seldom extending beyond classifying constituents in terms of subject, object and verb. No inspection of subjects’ propositional attitudes can account for their grammaticality judgements. Equally untenable is an account which invokes a stored mental repository of string-meaning pairings based upon the subjects’ past experience. Such a competence could not explain the fact that subjects’ grammaticality judgements are unbounded and very often novel. Because language allows for the iteration of attitudinal verbs and relative clauses,

---

<sup>53</sup> This is something of an oversimplification since natural languages are not constituted by well-formed formula in the way that formal languages are. Rather, the unacceptability of a sentence can be more or less severe. Nevertheless, I ignore this complication here.

<sup>54</sup> Chomsky has sometimes spoken of what speakers of a natural language must *know*. However, knowledge implies, if not entails, justificatory grounds, and insofar as Chomsky is not interested in making an epistemic claim, only an explanatory one (see, Chomsky, 1980: 69-70), the epistemically neutral term ‘competence’ is better in this respect (Collins, 2004; 2008: 125-8; Hinzen, 2006: 19).

speakers' grammaticality judgements extend to an unbounded number of strings, most of which they have had no past experience of:

5. The man which the dog bit which the cat scratched...
6. John thinks that Mary said that Bill hoped...

Prior to the inception of Generative Grammar, structural linguists tried to explain this novelty phenomenon by gesturing towards a capacity for analogy. For instance, it was said that a speaker "utters them [new sentences] on the analogy of similar forms which he has heard" (Bloomfield 1933: 275). However, there is simply no precise formulation of how an analogical ability could explain speakers' grammaticality judgements. Jackendoff (2002: 87) draws attention to the example:

7. Every acorn grew into an oak.
8. Every oak grew out of an acorn.
9. An oak grew out of every acorn.
10. \*An acorn grew into every oak.

From (7-9) there is a relation of symmetry between each acorn and each oak. However, in the case of (10) this relation is not available, despite its apparent analogical form. Similarly, Pietroski (2005: 3-5) underlines the non-analogical character of subjects' grammaticality judgements by citing what he calls 'negative facts'. Negative facts concern the unavailability of certain interpretations of strings. For example:

11. a. John is easy to please.  
       b. It is easy for us to please John.  
       c. It is easy for John to please us.
12. a. John is eager to please.  
       b. John is eager for us to please him.  
       c. John is eager that he pleases us.

(11. a) can be paraphrased with (11.b), but not with (11.c), whilst (12.a) can be paraphrased with (12.c), but not with (12.b). Pietroski points out that these unavailable interpretations, (11.c) and (12.b) are not incoherent or contradictory. Most interestingly from the current perspective is that were grammaticality judgements analogical, we would expect these interpretations to be available. For example, given that when 'John' combines with 'is easy to please', John is the recipient of the pleasing (i.e. 11.a. to 11.b), analogy would imply that when 'John' combines with 'is eager to please' John is the recipient of the pleasing also (i.e.

12.a to 12.b). Yet 12.b is not a possible reading of 12.a. Alternatively, given that when ‘John’ combines with ‘is eager to please’ John is the pleaser (i.e. 12.a to 12.c), analogy would imply that when ‘John’ combines with ‘is easy to please’, John is the pleaseer also (i.e. 11.a to 11.c). Yet 11.c is not a possible reading of 11.a. As Chomsky (1988: 20) points out, such examples preclude explaining subjects’ grammaticality judgements by recourse to a domain general capacity for reasoning analogically.

Since the late 1950s, linguists have argued that subjects’ grammaticality judgements are attributable to a computational competence: a generative grammar. This grammar is comprised of a finite number of lexical items along with rules and principles for recursively combining these items into part-whole structures. This grammar mechanically generates the grammaticality judgements of speakers, and so can be thought of as the causal antecedent of these judgements. For this reason, a generative grammar is commonly called ‘I-language’, since it defines the sentences of a speaker’s language intensionally, at the individual level, and in internalist terms (i.e. without recourse to social conventions).

### 3.3. *Moral judgements*

The primary explananda for moral psychologists are the moral judgements subjects make. In fact however, investigation is not only restricted in this way, but has been primarily limited more specifically to a certain *type* of moral judgement: deontic moral judgements. These judgements concern whether a particular act or event is permissible/obligatory/forbidden. This is a non-trivial limitation given that as well as deontic judgements, moral judgement also include evaluative judgements e.g. ‘Brian is good/courageous/kind’ etc.<sup>55</sup> Deontic moral judgements (from here on in, moral judgements) have been investigated as responses to situations of ‘quandry ethics’. Subjects are presented with complex vignettes describing a particular situation and a possible course of action. Subjects are then asked whether the proposed course of action is a permissible one. Consequently, the judgements elicited are of the form:

‘It is permissible for person *A* to do *X* in situation *S*’

---

<sup>55</sup> There are a number of reasons why the former have been favoured over the latter and I will only gesture at them here. Firstly, evaluative predicates are inherently gradable, prompting an issue of quantification which is hard, if not impossible to determine simply the basis of subjects’ responses. Secondly, deontic operators exhibit a common logic (Mikhail, 2007: 144). There is no such logic for evaluative terms, apparently foreclosing the possibility of a precise or formally specifiable relation between different evaluative predicates.

The vignettes most utilised have included the well known trolley problems (Mikhail, 2000, 2011; Hauser, 2006), originating in the work of Philippa Foot (1978). Alongside these, researchers have also devised their own scenarios of lifeboat dilemmas (Shenhav & Greene, 2010), infanticide dilemmas (Greene, et al., 2004) etc.<sup>56</sup> For example:

*Switch case:* A runaway trolley is about to run over and kill five people, but the driver can push a button that will turn the trolley onto a sidetrack where it will kill only one person. Is it permissible to push the button?

*Footbridge case:* A runaway trolley is about to run over and kill five people, but a bystander who is standing on a footbridge can shove a man in front of the trolley, saving the five people but killing the man. Is it permissible to shove the man?

The most important, general result of these studies has been to show that, as a matter of fact, peoples' moral judgements exhibit a certain regularity and predictability. For example, subjects overwhelmingly agree that it is permissible to push the button in the *switch case* and impermissible to push the man in the *footbridge case*, despite the equivalent consequences (five men saved, one killed). These kinds of regularities consistently appear for a wide range of such scenarios. Furthermore, this holds across demographic divisions such as sex, gender, religion and education (Mikhail, 2011; Hauser, 2006). Generalising from the data, we might say, rather sweepingly, that with respect to peoples' moral judgements, *as a matter of fact*, it is not the case that 'anything goes'.

This regularity thesis is especially significant when considered in conjunction with the *novelty* of subjects' moral judgements. As Mikhail (2000: 56-7) puts it, "with a few exceptions, no two situations which occasion moral judgements are exactly alike. Each is a potentially brand new combination of agents, acts, events and circumstances, occurring for the first time in the history of the universe". The trolley problems are a case in point. Few people will have ever been confronted by such peculiar situations, yet they have no difficulty in producing moral judgements about them. Quite simply, moral judgements are not restricted to a set of prescribed situations or ones which people have previously encountered. The vast majority of moral judgements subjects produce pertain to situations hitherto unfamiliar to them, involving novel combinations of actors and properties. Indeed, such considerations lead

---

<sup>56</sup> There are complex and nuanced issues surrounding the choice of dilemmas subjects are presented with. For example, how artificial should the scenarios be? How much information about the participants in the scenario should be supplied? (Participants are almost always described in the third person, hence foreclosing any personal connection to the subject).

Dwyer (2009: 10) to talk of the human capacity for an “infinity of scene/[deontic]valuations associations”. This novelty is important because it renders the *regularity* of subjects’ moral judgements all the more puzzling. In particular, it invalidates any explanatory account which simply attributes to subjects a stored repository of scene type/deontic valuation pairings. Such a competence could not account for the fact that moral judgements can be, and very often are, novel from the perspective of the subject.

The question for the moral psychologist then is why do people have the moral judgements that they have? More precisely, how can they be regular and predictable in the ways that they are, rather than regular in quite different ways, or indeed, not at all, and yet at the same time often be novel from the perspective of the subject? Answering this will be a matter of establishing “the causal etiology of such judgements” (Dwyer, 2009: 279) or reverse engineering their psychological origins. Despite the paradigmatically normative character of the judgements in question, tracing their genesis is, nevertheless, a purely descriptive undertaking and does not entail any normative conclusions.<sup>57</sup>

*Prima facie*, perhaps the most obvious source of these regularities is common patterns of reasoning and deliberation. Indeed, the rationalist thesis that peoples’ moral judgements are determined by the types of moral reasoning they engage in was fundamental to the highly influential model of moral development advanced by Lawrence Kohlberg (1981) (for discussion, Krebs & Denton, 2005: 629). However, one of the most significant developments in moral psychology over the past twenty years has been a recognition that moral judgement does not depend upon reasoning and deliberation in the way which Kohlberg envisioned. Of particular importance has been the phenomenon of ‘moral dumbfounding’, where subjects are unable to articulate reasons for their own moral judgements. Perhaps the most famous example of moral dumbfounding concerns sibling incest (Haidt, Koller and Dias, 1993), though the phenomenon has been exhibited across a wide range of cases (e.g. Cushman, et al., 2006). Moreover, where subjects do offer reasons, they are often not relevant, or clearly ill-suited to justifying the judgement made. Rather than being the products of deliberation, subjects’ judgements very often appear almost reflexive, appearing in consciousness as a ‘flash of insight’ (Haidt, 2001: 818). Indeed, Haidt claims that the field has converged upon

---

<sup>57</sup> Although there is no entailment, certainly some researchers have sought to forge a connection between descriptive facts about the genesis of a judgement and its normative warrant. “It is hard to see how one could reach a conclusion about whether moral intuitions [moral judgements] are justified without having any idea of how they work.” (Sinnott-Armstrong et al, 2010). Generally however, researchers have been quick to stress that an answer to the descriptive question, need not entail anything with respect to the normative one e.g. Hauser (2006: 96); Dwyer (2007: 252-253).

the ‘intuitive primacy principle’ which relegates verbal reasoning, in most cases, to the role of a post hoc rationalisation of a pre-existing intuition. According to this view, moral deliberation plays the role of a lawyer trying to defend a position already determined (Haidt, 2001: 814).<sup>58</sup>

The conjunction of (a) the regularity and predictability of subjects’ moral judgements, (b) their novelty and, (c) the lack of a clear explanatory connection between them and subjects’ moral reasoning, raises the prospects of a distinctive object of inquiry in the moral domain: that is, those psychological mechanisms and processes, those ‘secret springs and principles’, *which are* responsible for subjects’ moral judgements. This object of inquiry can be described in terms of Marr’s three levels. Consider first the functional level, where the challenge is to describe the function of the target mechanism in terms of its input and output. The output in the current case is clear: the moral judgements subjects produce. The far more challenging task is to explicate the input. For the language faculty this is a lexical array, roughly, a set of words, but what about in the moral case? An important initial observation is that whatever the input is, it is necessary that it encode for those distinctions and properties which the output provably co-varies with, and so those which *ex hypothesi* play a causal role in the derivation of moral judgements. *That* there are such distinctions is presupposed by the methodological approach universally embraced by moral psychologists: presenting subjects with descriptions of acts/events and then testing how their judgements systematically vary when features of the acts/events are modified. Indeed, insofar as these properties are generally supposed to be non-moral in character, moral psychologists are simply presupposing a descriptive version of moral supervenience:

“An action-guiding appraisal of actions, persons, and institutions that did not depend on the otherwise specifiable features of these items would not be moral judgement as we know it.” (van Roojen, 2006: 172)<sup>59</sup>

This seems reasonable enough. Even without a precise enumeration of what the “otherwise specifiable features” are, there is good evidence that they must include properties such as agent, patient, end, means, battery, omission etc., and consequently, the input must be such

---

<sup>58</sup> This is almost certainly an oversimplification (Pizarro and Bloom, 2003). Establishing the extent to which reason does shape moral judgements is a delicate, and largely open question. Nevertheless, for present purposes, the generalisation that moral judgements are typically not driven by deliberation suffices.

<sup>59</sup> Although moral supervenience is generally understood in the philosophical literature as a normative principle about how people *should* behave/judge, there no reason why it cannot also be understood as a descriptive principles about how people *do* behave/judge.



that it can encode or mark for them. However, while being reasonable enough, this premise actually has serious ramifications for what the input to functional system could be:

“...an intervening step between stimulus and response must be postulated: a pattern of organisation of some sort that is imposed on the stimulus by the mind itself” (Mikhail, 2000: 130)

Subjects generally judge it permissible to flip the switch in the *switch case*, but not push the man in the *footbridge case*, despite the equivalent consequences, and a plausible explanation of this invokes an ‘intentionality difference’: the killing of the man in the *footbridge case* is intended, whereas the killing in the *switch case* is merely foreseen. Yet, this ‘intentionality difference’ is not explicit in the vignettes subjects are presented with and short of linguistically specifying this intentionality, it is hard to see how this difference *could be* explicit. This point applies generally to other properties and distinctions which subjects’ considered moral judgements are sensitive to: including ‘means’, ‘ends’, ‘act’, ‘omission’ and, according to Mallon and Nichols (2010: 307), ‘right’ and ‘injury’. Such properties are not fixed or determined by the stimulus given to subjects, whether that be the artificial vignettes of a research lab or visually perceived in the ‘real world’, but rather, need to be in some sense inferred from such stimulus. Indeed, it is unclear whether it makes any more sense to attribute these distinctions to the nature of the stimulus than it does to attribute grammatical categories to acoustic segments. Just as an acoustic stream needs to be parsed by a subject according to grammatical categories e.g. noun phrase and verb phrase, so too these distinctions relevant to the moral valuation should be understood as *forms of parsing*. In fact, insofar as it is possible to morally evaluate imagined or counterfactual situations there need not be any stimulus at all (Sterelney, 2010: 287). For example, consider Gray et al.’s (2012: 103) claim that, “there need not “objectively” be an intentional agent and a suffering patient in every moral situation, but only the perception of this dyad”.

In other words, we need to distinguish the stimulus which may trigger a moral judgement, from the input which is psychologically mapped moral judgement. It should be clear that this is no more than a morality-specific formulation of the under-determination thesis noted in Section 1.5. There is a gap between *distal* stimulus and *proximal* stimulus and consequently, “psychological regularities will always have to be stated relative to particular

readings of stimulus” (Pylyshyn, 1980: 11).<sup>60</sup> Echoing Pylyshyn then, we can say that in the moral domain, the relevant distal stimulus must be parsed as an internal representation. Call this parsing the construction of an *act analysis*:

“...the brain must be generating actions representations of its own that go beyond the information given. That is, much like a given patch of retinal stimulation or the acoustic stream in speech perception, the stimulus here evidently consists merely of clues for the formation of an unconscious percept that the perceiver firsts constructs using her own internal resources.” (Mikhail, 2011: 114)

The position here is that, adopting a representational theory of mind, subjects utilise various stimuli drawn from different sensory modalities to construct unconscious, act analyses. These act analyses are complex mental representations and constitute the ‘input’ to the system described at the functional level. Although this leaves many substantive theoretical questions about these constructions open, it is widely agreed that some form of act analysis must be postulated by any empirically adequate account of moral cognition (Cushman, Young and Greene, 2010: 6-9; Prinz, 2008: 161-2).<sup>61</sup>

The second of Marr’s levels of description is the algorithmic level: according to what rules is the input mapped to output? As Joseph and Haidt (2008: 379-380) explain, the challenge at this level of analysis is to describe “input-output programming connecting the pattern in the social world... to an evaluation”, or alternatively, “formulating a set of algorithms from which they [the moral judgements subjects produce] can be derived” (Mikhail, 2009: 28).<sup>62</sup>

---

<sup>60</sup> Although in the previous section I chose not to discuss how this under determination thesis applies to grammaticality judgements, there clearly is a linguistic analogue: the physical marks a speaker is confronted with are often ambiguous and can be assigned multiple interpretations e.g. ‘Flying planes can be dangerous’ can be parsed as either [<sub>VP</sub>[<sub>V</sub> Flying][<sub>NP</sub> planes] can be dangerous] or [[<sub>NP</sub>[<sub>Adj</sub> Flying][<sub>N</sub> planes]] can be dangerous]. In such a case, there is a distinction between the distal stimulus (the marks on paper or acoustics) and the proximal stimulus, where the former fails to fix or determine the latter.

<sup>61</sup> Mikhail (2011: Chpt. 5) has proposed what is, to date, the most developed account of these representations. Although I have left the matter fairly open here, I think that understanding the nature of act analysis is integral to understanding moral cognition as a whole. For example, there is good reason to think that an act analysis must be a part-whole organisation of atomistic constituents. Another issue is whether act analysis should be understood as unique to the moral domain (see Sterelney, 2010; Dupoux and Jacob, 2007). It is quite plausible that the act analysis necessary for moral judgement is isomorphic with that found in event cognition generally (Strickland, et al., 2012), raising interesting questions.

<sup>62</sup> A shortcoming of this way of posing the challenge is that the type of empirical adequacy being targeted is merely extensional rather than intensional (Nichols, 2005). Nevertheless, extensional adequacy is challenging enough to be getting on with. In order to penetrate deeper to intensional adequacy, if it is methodologically possible at all, acquisition considerations will presumably be key – as they have been within Generative Grammar.

Figure 3.



These rules are properly understood as *functions* which, operating on act analyses, yield deontic valuations as their range; these functions can be paraphrased as follows: if an act analysis has feature  $X_1, X_2... X_n$ , map it to the deontic value,  $\alpha$ . Of course, this is a great oversimplification insofar as individual features must be treated as having *pro tanto* rather than sufficient force, but the general point stands. The hope is that these rules will prove amenable to explicit formulation, in the same way that the grammatical rules of language users have proven to be (Harman, 2008: 7-8).

Whether an empirically adequate set of rules would be unique to the moral domain is a keenly contested issue in the literature. Some commentators argue that there is good evidence for principles and rules unique to the moral domain and that these may constitute a moral faculty or moral grammar.<sup>63</sup> This is by no means universally accepted though, and it has been argued that domain general rules may suffice. For example, calculations of utility magnitude and the probability of success have been shown to play a role in determining the moral appraisals subjects make in at least some situations (Shenhav and Greene, 2010). The following principles have been postulated in order to account for subjects' moral judgements (though whether they are unique to the moral domain I set aside here):

*The Action Principle:* Harm caused by action is morally worse than equivalent harm caused by omission.

*The Intention Principle:* Harm intended as a means to a goal is worse than equivalent harm foreseen as the side effect of a goal

*The Contact Principle:* Using physical contact to cause harm to a victim is worse than causing equivalent harm to a victim without using physical contact (Cushman, Young, Hauser, 2006).

<sup>63</sup> See Kirkby (2013); Kirkby, Hinzen & Mikhail (2013); Cushman, Young and Hauser (2006); Dwyer (2007, 2009); Harman, (2008); Hauser (2006); Hauser, Young and Cushman (2008); Mikhail (2000, 2007, 2011) .

*The Principle of Double Effect:* An otherwise prohibited act which has good and bad effects may be permissible if the act itself is not directly intended, the good but not the bad effects are directly intended, the good effects outweigh the bad, and no morally preferable alternative is available.

## Part II

# KANT: THE INTENTIONALITY OF JUDGEMENT

That Kant's transcendental philosophy contains or amounts to a reaction to Hume is hardly an original observation. Kant himself affirmed as much with his famous remark that it was Hume who awoke him from his "dogmatic slumbers" (Prol: intro., 6). In Strawson's *The Bounds of Sense*, the nature of this reaction is carefully detailed. Strawson argues that, in contestation of Hume's scepticism, Kant shows that the existence of the external world is a necessary condition for self-consciousness. Since sceptics generally accept some notion of self-consciousness, they must also accept the existence of the external world. However, irrespective of whether this is a *good* interpretation of Kant, it would hardly be a *relevant* one in the context of this thesis, given that the previous chapter was devoted to emphasising the primacy of Hume's naturalistic or constructivist outlook.<sup>64</sup>

Instead, the 'Kantian reaction' I am interested in here concerns *intentionality*. Kant thinks that there is a problem regarding the possibility of intentionality: a problem which strikes at the heart of what it is to be a being which *cognizes* or enjoys *experience*: a problem which Hume, in his best naturalist clothing, had failed to adequately address. In Chapter four, I describe the nature of this intentionality problem, endeavouring to situate it within Kant's overarching transcendental philosophy. Chapter five concerns Kant's philosophical response to this problem: his account of how intentionality is possible. There are two essential and interdependent components of this account: that of rules and that of judgement. As such, Kant's theory of judgement is an integral part of his *solution* to the intentionality problem. Chapter six describes how the central role allocated to judgement leads Kant to ascribe special significance to the particular forms which judgements exhibit. However, I argue that

---

<sup>64</sup> I do not mean to imply that the suggestion that Kant responds to Hume-as-naturalist, rather than merely Hume-as-sceptic, has gone unnoticed in the secondary literature. On the other hand, this critique has not been nearly as prominent as the Strawsonian line and is seldom developed in any great detail.

Kant does not go far enough, and actually *understates* their significance. In particular, I contend that the distinction Kant draws between the ‘merely’ formal forms of judgement and the categories is not sustainable. Rather, the forms of judgement are of inherent transcendental significance. Finally, Chapter Seven addresses the issue of how the forms of judgement are to be identified. I argue that these forms ought to be derived from our best theories of *grammar*. That is, grammar provides the table of judgement forms. This amounts to a naturalisation of this component of Kant’s philosophy.

## Chapter 4. The Nature of Experience

---

### 4.1. From synthetic a priori judgements to the possibility of experience

The *Critique of Pure Reason* (CPR) begins on a rather mournful note. Metaphysics or speculative philosophy, Kant laments, is in a parlous state, plagued by dogmatism on the one hand and on the other, scepticism. The problem, according to Kant, revolved around the uncertain epistemic status of a class of judgements he called ‘synthetic a priori’. For example:

1. The shortest distance between two points is a straight line.
2. All succession is alteration.<sup>65</sup>
3. Every event has a cause.

In calling these judgements ‘synthetic a priori’ Kant means, in brief, that (a) they are not truths of logic and that, (b) they cannot be known by recourse to experience. As such, they amount to a ‘third type’ of judgement, defined in contradistinction to those judgements which, on the one hand are truths of logic, and on the other, those judgements which are known by recourse to experience. Consider the former first:

4. Bachelors are unmarried men

(4) is an example of an *analytic* judgement for Kant: it is a logical truth. This might appear surprising – it doesn’t seem to be the kind of truth which logic yields. However, crucial to this claim is Kant’s view that concepts can be decomposed into constituent marks.<sup>66</sup> For example, according to Kant, the concepts ‘unmarried’ and ‘man’ are contained under that of ‘bachelor’ as marks. A perspicuous analysis of (4) then, would be (5):

5. For any  $x$  to which bachelor ( $\text{unmarried} \wedge \text{man}$ ) applies, unmarried man applies.

---

<sup>65</sup> That is, alteration of a persisting substance.

<sup>66</sup> With the publication of Quine’s ‘Two Dogmas of Empiricism’ (1951), in which the idea of ‘concept containment’ was subjected to scathing criticism, the credibility of a category of analytic judgements fell into disrepute. Despite this, it is not clear to me that there is anything tendentious about the notion of analyticity (for a nice discussion, see Bell, 2001: 9). For one thing, the theory of concepts it is founded upon has a comparable modern analogue in lexical semantics (cf. Jackendoff, 2002: Chpt. 11). Either way, in the context of this thesis, nothing hangs on the philosophical credentials of the analyticity which I am here using primarily a point of entry to other topics in Kant’s work. In principle, these topics could be formulated without presupposing anything about analyticity.

This analysis reveals that (4) has analytic content: the predicate is ‘contained’ in the subject concept and so the content can be formulated in tautologous terms i.e. ‘For any  $x$  to which  $P_1 \wedge P_2 \dots$  applies,  $P_1$  applies’.<sup>67</sup> As such, the truth of an analytic judgement is grounded the law of non-contradiction. Despite their logical triviality, analytic judgements are not utterly uninformative or useless. They can clarify and inform us of the nature of our concepts, many of which may be quite opaque to us. Nevertheless, because the truth of analytic judgements is secured irrespective of how the world is, i.e. merely by the laws of logic, they do not extend our knowledge of the world. For example, the truth of (4) is consistent with the state of the world turning out to be any of an infinite number of possible ways: realising any set of facts, and consequently its truth cannot inform us of how the world is. For this reason, Kant says that analytic judgements are non-ampliative. By contrast, consider:

- 6. Some tables are round
- 7. Humans inhabit cities

These judgements are, evidently, not analytic. No plausible analysis of the concepts ‘table’ and ‘round’ will show that the predicate is contained in the subject: they cannot be formalised in tautologous terms and so are not truths of logic. Kant calls judgements which are not truths of logic ‘synthetic’. Insofar as they make a claim which goes beyond the law of non-contradiction, they are ampliative. As we might say, they make a substantive claim about how the world is. Whereas, analytic judgements are justified by the laws of logic, most synthetic judgements are justified by recourse to empirical experience. For example, ‘some tables are round’ is known to be true because we have had empirical experience of round tables. Judgements which are justified on the basis of experience are labelled ‘a posteriori’. So both (6) and (7) are synthetic a posteriori judgements.

There is an important point which needs to be made about a posteriori judgements, or judgements which are justified by recourse to empirical experience. According to Kant, a posteriori judgements are only capable of expressing contingent, as opposed to necessary, truths. This is because, “experience does indeed teach us that something is thus or thus, but not that it cannot be otherwise” (B4). For example, insofar as (6) and (7) are both justified by our empirical experience of the world, it is clear that they can only be expressing contingent

---

<sup>67</sup> An objection to this formulation of analyticity is that it is restricted to categorical propositions (Quine, 1951: 21). However, it can be extended to other proposition types without difficulty. For example, a conditional would be analytic if, firstly, the antecedent and consequent have the same subject and, secondly, if the predicate in the consequent is a mark of the predicate in the antecedent.



truths about how the world happens to be, not about how the world must be. A posteriori judgements therefore, are always contingently true, if true at all. In this respect, Kant agrees with Hume: the deliverances of the senses are never necessities.

The category of synthetic a priori judgements lies outside of either of these two classes of judgements. In contrast to analytic judgements, Kant considers that (1), (2) and (3) all make substantive, ampliative claims about how the world is: claims which cannot be reduced to, or justified by, the laws of logic. Equally however, (1), (2), and (3) are also said to be knowable a priori, setting them apart from synthetic posteriori judgements. Specifically, Kant thinks that they amount to claims of necessity, rather than mere contingency. In outlining this class of synthetic a priori judgements, Kant saw himself as delineating the subject matter of speculative philosophy, or metaphysics. Metaphysical inquiry is a priori by nature, and it is concerned with synthetic judgements insofar as it seeks to advance our knowledge of objects or the world, rather than simply clarify the nature of our ideas. According to Kant, synthetic a priori judgements are the appropriate subject matter for metaphysics.

“...on such synthetic, i.e. expansive, principles depends the whole final aim of our speculative a priori cognition” (A10/B14)

“...metaphysics consists... of nothing but synthetic a priori propositions” (B18)

The difficulty with the synthetic a priori however, and therefore the difficulty with metaphysics, concerns how such judgements can be justified. In this respect, synthetic a priori judgements are problematic in a way in which analytic or a posteriori judgements are not. Analytic judgements are justified by the rules of logic. A posteriori judgements are justified by experience. But since Kant defines synthetic a priori judgements as being neither truths of logic, nor justifiable by recourse to experience, what does justify them?

This problem can be described as that of how it can be legitimate to apply concepts to objects of experience in ways not justified by either logic or empirical experience (Winkler, 2010: 58). For example, in the judgement ‘every event has a cause’, the concept of having a cause is being applied to every  $x$  which satisfies the concept of event. The question is, what justifies this application of the concept? *Ex hypothesi*, it is not justified by experience or by logic, but then what does justify this application? Crucially, Kant’s problem here is an epistemic one, rather than a psychological one. His aim not to, for example, challenge

Hume's psychological account of how we form beliefs; nor is it to say that we must have innate ideas of one form or another. Rather, the question is one of entitlement (*quid juris*) (A84/B116).

In his earlier work, e.g. the *Inaugural Dissertation*, Kant thought that this problem could be resolved by appeal to the real use of the intellect: the nature of mind-independent objects could be apprehended by the intellect without the aid of the senses. By the time of the CPR however, he had abandoned this position (Wolff, 1963: 15-17). As Kant explains in his letter to Herz, in espousing a real use of the intellect, he had "passed silently over" a crucial problem, that of, "how a representation that refers to an object without in any way being affected by it can be possible" (Corr: Letter to Herz, Feb. 1772, 133). The short answer to this, Kant decides, is that it is not possible; there is no real use of the intellect and any claim to the contrary is mere dogmatism. Instead, "all knowledge begins with experience" (A1).<sup>68</sup> In this respect at least, Kant was far closer to his empiricist predecessors than his rationalist predecessors. This stance entailed that, of the 'supersensible' objects which had occupied rationalist metaphysicians e.g. God, the soul, objects which, by definition, transcend the bounds of possible experience, no knowledge was possible (Pippin, 1982: 10).<sup>69</sup>

However, this meant that Kant was mired in a severe predicament. The whole field of metaphysics depends upon showing that synthetic a priori judgements are possible (that they can be legitimate), yet his philosophical commitments barred him from appealing to either the real use of the intellect or empirical experience of mind independent objects: a disjunction which apparently exhausts the possibilities. At this point in the dialectic, Kant makes a simple, yet philosophically revolutionary suggestion:

"...it has been assumed that all our cognition must conform to objects. On this presupposition, however, all our attempts to establish something about them a priori, by means of concepts through which our cognition would be expanded, have come to nothing. Let us, therefore, try to find out by experiment whether we shall not make better progress in the problems of metaphysics if we assume that objects must conform to our cognition." (Bxvi)

This is, of course, Kant's famous Copernican turn. By talking of objects *conforming* to our cognition, Kant is invoking the idea that cognition presupposes certain conditions: a priori

---

<sup>68</sup> 'Experience' should be interpreted here as essentially containing a sensory or receptive component.

<sup>69</sup> Although of course, this does not preclude these concepts playing a regulative role in our reasoning.

conditions, since it is only by conforming to these that objects can be cognised in the first place. Kant declares that he will call ‘transcendental’ that which “deals not so much with objects as rather with our way of cognizing objects in general insofar as that way of cognizing is to be possible a priori” (A11-12/B25). That is, the transcendental is concerned with the a priori conditions for the possibility of cognition.

It is crucial to emphasise that Kant portrays the Copernican turn as a philosophical *undertaking*, rather than a substantive view or position in its own right. It is referred to it as an “experiment” (Bxvi), and the reason for this is obvious. Quite simply, at this stage, the existence of a priori conditions of cognition is completely unproven. Nothing has been said which indicates, let alone justifies the postulation of such conditions. Whether there are such conditions is precisely what is in question, and as such, we might follow Hanna (2001) in referring to the Copernican turn as the adoption of a transcendental *method*: one which seeks to uncover a priori conditions of cognition.

There are, I think, two related issues regarding this transcendental method which need to be addressed from the outset. Firstly, what motivates this undertaking? Call this the motivation issue. As Longuenesse (1998: 7 & 71) emphasises, in proposing the Copernican turn, Kant moves from the question:

- How are synthetic a priori judgements possible?

to

- Under what conditions is cognition possible a priori?

Although the existence of such a priori conditions is not, at this point, proven, clearly Kant thinks that establishing this will bear upon the problem of the synthetic a priori. Were there no such connection between these two questions, it would be completely obscure as to why we have moved so abruptly from one to the other. Kant quickly reassures us on this point when, almost immediately after announcing the Copernican turn, he declares that if there are a priori conditions of cognition, he can “quite readily conceive” how we might have synthetic a priori knowledge (Bxvii). The claim is that establishing the existence of a priori conditions of cognition will somehow ‘solve’ the problem of the synthetic a priori, and it is this connection which motivates the transcendental method. This much is clear. However, what is *not* yet clear is exactly what this connection is for Kant. Certainly, Kant thinks that there is such a connection, but equally certain is that we are not yet in a position to state precisely

what this connection amounts to. We are still missing a piece of the puzzle. In order to see what this piece is, we must turn to the second issue regarding the Copernican turn which needs to be addressed.

Suppose we want to pursue the transcendental method, adopt the Copernican turn and seek out conditions for the possibility of cognition. Where do we start? Well, the point of departure for any such inquiry will be, surely, *that for which* a priori conditions are being sought: namely, cognition or *experience*.<sup>70</sup> It is senseless to inquire after conditions of experience, in the absence of a clear conception of what experience *is*. In other words, the point of departure for the application of the transcendental method is a robust answer to the question:

- What is experience?

This is the ‘second issue’ regarding the Copernican turn: as a methodological approach, it presupposes a well defined view of what experience is. This is the *premise* from which inquiry must proceed, for in its absence, talking of the a priori conditions of experience verges upon vacuity. In the next section, I outline what this premise is for Kant: what he means by ‘experience’.

Returning to the motivation issue, this illuminates why we are not yet in a position to explain the connection between the synthetic a priori and the conditions of experience: the piece of the puzzle which is missing *precisely is* this conception of experience – something we are currently lacking. Consider the following:

- (i) What is experience?
- (ii) Under what conditions is experience possible a priori?
- (iii) How are synthetic a priori judgements possible?

Kant’s view is that there is a crucial connection between (ii) and (iii): an answer to (ii) will bear upon answers to (iii). However, as I have shown, (ii) is not even meaningful until we have an answer to (i). As such, any purported connection between (ii) and (iii) is dependent upon first addressing (i). In light of this, I shall give a full account of this connection between the synthetic a priori and the a priori conditions of cognition only in the final part of the next section, once the nature of experience has been suitably clarified.

---

<sup>70</sup> It is common in Kantian scholarship to use the terms ‘cognition’ and ‘experience’ interchangeably, and I shall be following this practice here.

## 4.2. Relation to an object

The notion of experience is the fundamental premise of the transcendental method. It is the centrepiece around which all transcendental inquiry is organised. What then, does Kant mean by experience?

“I noticed that I still lacked something essential, something that in my long metaphysical studies I, as well as others, had failed to pay attention to and that, in fact, constitutes the key to the whole secret of hitherto still obscure metaphysics. I asked myself: What is the ground of the reference of that in us which we call “representation” to the object?” (Corr: Letter to Herz, Feb. 1772, 132-133)

“The a priori conditions of a possible experience in general are at the same time conditions of the possibility of objects of experience.” (A111)

Kant persistently invokes ‘relation to an object’ as the mark of experience.<sup>71</sup> The concept of an object, in Kant’s hands, is not an ontological category; it is not supposed to mark a difference between, for example, desks and stars on the one hand and the greenness of a leaf or Will’s love for Kate on the other. Rather, ‘object’ encompasses all beings including, for example, events (Bird, 1962: 89), not merely some subset thereof. As such, the relation of representations to an object is equivalent to ‘objectivity’ or ‘objective experience’ or ‘objective validity’ all of which Kant uses interchangeably in the CPR, and I shall follow him in this respect.<sup>72</sup>

---

<sup>71</sup> It might be objected at this point that it is *the unity of apperception* which is the fundamental premise of the transcendental method. This is certainly the traditional view (Bennett, 1966; Strawson, 1966: 26; Wolff, 1963: 105) and indeed, Kant himself calls it the “supreme principle in all of human cognition” (B135). It is crucial to stress therefore, that my taking the relation of representations to an object as the premise is not in contradiction of this view. For Kant, they are equivalent:

“The essential point is that thinking its own identity is not another act that the mind performs in addition to the representation to itself of objective connection.” (Allison, 1996: 61)

Given that there is no contradiction here, I think restricting myself to the ‘object side’ of (what is for Kant) a single object/apperception coin, is well motivated. Firstly, I am not attempting a complete exegesis of CPR here – apperception is one of many central strands which will not receive direct attention. Secondly, while they *are* equivalent for Kant, there is nothing non-trivial about this claim in philosophical terms, and since, in the context of this thesis what matters is the issue of the representation of the object, I do not want to render it more tendentious by asserting its equivalence to apperception needlessly.

<sup>72</sup> Admittedly, Kant’s use of the term ‘object’ is not always consistent. For example, “Everything, every representation even insofar as we are conscious of it, may be entitled object. But it is a question for deeper enquiry what the word ‘object’ ought to signify in respect of appearances when these are viewed not insofar as they are (as representations) objects, but only insofar as they stand for an object.” (A189-90/B123-5)

Furthermore, objective experience is not true or veridical experience, or at least, it need not be. Crucially, while it is natural to interpret Kant's 'object talk' it as marking the difference between experience merely deemed to be true by a subject, and cognition which is actually true *of the object*, this would be a mistake. As Ameriks (1982: 12) points out, experience need not be true; experience can be both objective and delusive. For instance, perceiving a stick as bent when in water would be an example of objective cognition for Kant, despite the illusory nature of the experience. As Kant says:

“...a cognition is false if it does not agree *with the object to which it is related*”

(A58/B82-3; see also, Meta: 233-234) [my emphasis]

“The difference between truth and dreaming is not ascertained by the nature of the representations which are referred to objects (for they are the same in both cases).”

(Prol: §13, 37)

In other words, a representation being related to an object is no indication of its veridicality (Allison, 2004: 88; Longuenesse, 1998: 82), and therefore, while Kantian experience has occasionally been defined epistemically as *empirical knowledge* (e.g. Wolff, 1963: 98) this is, at best, misleading. Knowledge presupposes truth: objective experience does not.<sup>73</sup> Insofar as the transcendental method concerns experience then, it does not concern, or is insensitive to the difference between appearance and reality, illusion and veridicality (Bird, 1962: 159). Whether an object exists or a certain state of affairs realised is an empirical, not a transcendental matter (A376-7; Allison, 2008: 109). Given this though, what *does* Kant mean by 'relation to an object'?

“Kant's concern is not knowledge so much as the directedness of thought at objects, the intentionality or objective purport, that is a prerequisite for anything to be even a candidate to be a case of knowledge.” (McDowell, 2009: 210)

“Kant takes as his initial focus *intentionality* rather than *knowledge*.” (Brandom, 2002: 23)

In Brandom's terms (2009: 29), in speaking of 'relation to an object' Kant is targeting representational *purport*, and bracketing the further epistemic matter of representational *success* or *accuracy* (see also, Kukla 2011: 9-10). By 'relation to an object' Kant means to

---

<sup>73</sup> Hence, I think that Ameriks (1978) is seriously mistaken to argue that empirical knowledge is the premise from which the transcendental deduction 'regressively proceeds'.

denote the respect in which representations or mental states are *representational*: the sense in which they are *about* or *directed* at the world in some particular fashion: the respect in which they point beyond themselves. As Brandom (2009: 29) puts it, Kant is targeting what it is to:

“take or treat them [representations] as, for them to show up to us as, *representings*, in the sense of something that answers for its correctness to what thereby counts as being represented... This issue is the core around which cluster other elements of Kant’s concern with what he calls “objectivity”.”

It should be clear that hallucinations and non-veridical states exhibit objectivity of this kind. They purport to represent the world as being a certain way; for if they did not, they could not be non-veridical. It follows from this that ‘relation to an object’ is not really *relational*, if by that we mean that whether it is realised or not depends upon how the world turns out to be. Rather, objectivity is best thought of as being inherent to certain mental states (Bennett, 1966: 127), or as Allison (2008:108) puts it is “immanent to consciousness”. In contemporary philosophical parlance, this representational purport is known as *intentionality*.<sup>74</sup> That in virtue of which a state or representation is directed at a particular object is known as its intentional or semantic *content*. Crucially, the content of an intentional state is not the object towards which it is directed, but rather, that property which determines that it is so directed. On this view then, Kant’s concept of *experience* is semantic, rather than metaphysical or epistemological (Landy, 2009: 1; Hanna, 2001: 16).

There are two distinctive qualities of intentionality for Kant. The first is existence independence (Pereboom, 1988). This has already been gestured at by the separation of representational purport from representational success. To be in an intentional state is to represent things as being ‘such and such’ *irrespective* of their being so represented. Experience is ‘objective’ in the sense of being directed at or about a realm which can be distinguished from, and exists independently of, the fact of it being represented, the fact of experience itself.<sup>75</sup> The relevant notion of an *object* therefore is

“an object corresponding to and therefore *distinct* from cognition” (A104) [my emphasis]

<sup>74</sup> Guyer (1987: 79-80) argues that Kant equivocates between a concept of experience which would be accepted as a premise by empiricists, but not sceptics, and one which would be accepted by sceptics. However, insofar as this premise relies only upon the notion of representational purport, and brackets the question of representational success, I think that this would be acceptable to both philosophical camps.

<sup>75</sup> This is slightly misleading insofar as my own mental states can become the objects of other intentional states, but I set this point aside here.

Importantly, this existence independence requirement is equivalent to the possibility of error (A293). If intentionality presupposes that one be able to existentially distinguish the objects at which one's intentional states are directed from these states themselves, it also presupposes that one recognises that the objects at which intentional states are directed may not be as they are represented to be. That is, because the objects of intentional states are recognised as not being, in any relevant sense, constituted by their being so represented, but as being as they are irrespective of this, the objects are recognised as possibly being other than how they are represented (Davidson, 2004). As Hanna (2001: 17) points out then, the notion of 'object' relevant to intentional states for Kant is comparable to Brentano's 'intentional object': that upon which intentional states are directed, even if it does not exist.

The second central property of intentionality for Kant is that it is concept dependent or intensional. Intensionality is exhibited where the objects of intentional states are grasped by subjects under characterisations or concepts which are more finely-grained than the objects themselves. For example, there is a difference between representing that the cat went up an oak tree, and representing that the cat went up the oldest tree in sight, *even if* the oldest tree in sight is an oak tree (Davidson, 1982: 320). Representing the cat as going up the oldest tree in sight is not the same as representing the cat as going up an oak tree, even where the oldest tree in sight is an oak tree for the simple reason that the subject may not be aware of this fact. This shows that these are intensional states. Equivalently, a state is intensional if, in the attribution of the state to a subject, co-extensional terms are not substitutional *salva veritate*. For example, 'S represents that the cat went up the oldest tree in sight' may be true, yet 'S represents that the cat went up an oak tree' may be false, despite the only change being the substitution of co-referential terms. What intensionality indicates is that the relevant state is concept dependent. That is, objects are not picked out 'directly', but via the mediation of concepts, which *ipso facto* are more fine-grained than the objects they denote since any object satisfies numbers of conceptual descriptions. If a state is intensional, the intentional content of the state – that in virtue of which a relation to a particular object is established – is conceptual or descriptive.

Intensionality or concept dependence is a phenomenon which has been widely recognised to be integral to intentionality (Chisholm, 1957). Similarly, the notion of intentionality Kant is interested in is thoroughly intensional (Pereboom, 1988; Howell, 1992: Chpt. 5). All experience is said to require the use of concept (A106), introducing a descriptive element to all intentional states. Furthermore, Kant's logic is destitute of singular terms



(A70/B95; see, Anderson, 2008: 103), which are the best candidates of representations which relate to objects, ‘directly’. For Kant, all relations to an object are mediated by conceptual description.

This conception of intentionality defines experience for Kant. As such, it serves as the premise from which Kant’s transcendental method departs. Specifically, the premise is that that there are mental states which are experiential *in this way*. This is something Kant takes for granted; he does not seek to justify it directly. It is, therefore, both the premise and the presupposition of the investigation. That certain mental states are intentional or bear a relation to an object is foundational for Kant. I do not think there is anything objectionable about such a foundation. As Stevenson (1979: 348) says:

“...the critical [transcendental] philosophy must... appeal at bottom to certain natural facts about human beings; facts, however, which are so obvious and all-pervasive that they need no confirmation by empirical research”

I broadly concur with Stevenson, but two qualifications are in order. Firstly, labelling the relevant facts ‘natural’, as Stevenson does, is either vacuous or tendentious and as such, should be dropped. Secondly, the idea that these facts are ‘obvious’ or all-pervasive’, while plausible, should not be overstated. After all, philosophy is aimed at calling what is obvious into question. One could be an eliminativist about intentionality, so it cannot be *that* obvious; it is not incontrovertible. Nevertheless, in taking intentional experience as a premise, Kant can be understood as proceeding in a comparable manner to contemporary philosophers of mind. Roughly, just as such philosophers make claims about the nature of the mind e.g. that at least some mental states exhibit phenomenal consciousness, and just as some of these claims are ‘foundational’ in the sense that they constitute bedrock upon which debate and discussion is founded, so too Kant is departing from a premise about how beings like us are minded: specifically, they enjoy intentional mental states. This is a fact which Kant considers to lie at the heart of the human condition itself, as being fundamentally bound up with the kind of beings that we are (A546-7/B574-5).<sup>76</sup> Although this is not incontrovertible, the claim that at least some mental states are marked by intentionality is, to this day, one of the most

---

<sup>76</sup> This is clearly incompatible with Strawson’s (1966) ‘analytic’ reading of Kant’s transcendental deduction, whereby the premise from which investigation proceeds is that of the mere *idea* or *concept* of an experience we can make intelligible to ourselves. By contrast, my view is that Kant is interested in “conditions of our having the type of experience that we do” (Stapleford, 2008: 6). Though I will not undertake a critique of Strawson’s reading here (in my view, this has been sufficiently covered elsewhere, e.g. Kitcher, 1990; Hanna, 2001), it is worth noting that Strawson’s reading is incompatible with Kant’s clear belief that the discovery of a priori conditions of experience resolves the problem of the synthetic a priori.

secure (if not *the* most secure) of philosophical theses regarding the nature of the mind. Hence, the ground Kant is standing on is, arguably, as firm as anything available.

With the premise of the transcendental method in hand, it is possible to clarify the nature of the investigation to be undertaken. Specifically, to inquire after the a priori conditions of experience, is to inquire after the a priori conditions of *intentional cognition*. Brandom (2002: 22-23) describes the matter well:

“Cartesian skeptic asks what reason we have to suppose that the world is as we represent it to be in thought. An inquiry into the conditions of *successful* representation is accordingly an appropriate road to a response. Kant takes as his initial focus *intentionality* rather than *knowledge*. He asks about the conditions of even *purported* representation. What makes it that our ideas so much as *seem* to point beyond themselves, to something that they are *about*?”

This could be variously described as Kant’s ‘question’, ‘problem’ or ‘challenge’. In any case, it defines his transcendental undertaking. There are three points about this question which deserve attention. Firstly, in the context of this thesis this problem is especially significant since, irrespective of its centrality to understanding Kant, it also serves as a firm point of contact between Kant’s transcendental investigation and Hume’s science of man. Specifically, inquiring after the possibility of objective cognition is precisely the question Hume himself grapples with in the form of: ‘how do we come to believe in the existence of external bodies?’ In other words, insofar as ‘experience’ is distinguished by intentionality, and insofar as Kant’s transcendental method seeks conditions for the possibility of experience, we have arrived at the intentionality problem outlined in Section 2.3. Allison (2008: 208) is quite wrong therefore, to suggest that a possible Humean response to Kant’s ‘objectivity talk’ could be, “simply [to] deny that there is any objective validity as Kant understands it”. On the contrary, Kant’s question is a question for Hume also, albeit framed within a much richer philosophical context. This establishes a strong point of contact between Hume’s science of man and Kant’s transcendental philosophy, even a certain degree of intellectual continuity. Kant is essentially picking up a question which, I have argued, Hume tries and fails to adequately address.

The second point regarding Kant’s question is that it is not a *semantic* question. True, as I have noted, the premise upon which the investigation is founded, or from which the deduction proceeds, is a semantic one. Yet some commentators have gone further, apparently

suggesting that not only is the premise of the deduction a semantic one, but that the nature of the investigation itself is semantic.

“If I am correct, then the overarching purpose of the first *Critique* is to explain how a mental representation can refer to its object. This is the *Semantic Problem*.” (Hanna, 2001: 67; see also, Landy, 2009)

In my view, this is pretty clearly a mistake, or at the very least, misleading. Kant’s question, ‘what are the a priori conditions of experience’, amounts to the question of how semantic content *tout court* is possible; it targets the very possibility of semantics. This question has no analogue in contemporary philosophical semantics, which principally centres upon two broad issues. The first issue concerns how representational items, e.g. expressions, are about what they are about, or equivalently, what the semantic content of particular expressions or classes of expressions is. For example, are proper names referential or descriptive? The second issue concerns semantic composition. Roughly, given certain semantic contents, how do they combine so as to yield composite semantic contents? For example, how does the semantic content ‘dogs’ compose with the semantic content ‘swim’ to yield the complex semantic content ‘dogs swim’? Kant’s question however, is distinct from either of these issues. While philosophical semantics, we might say, works ‘within’ the basic paradigm of semantic content, Kant is working outside of it in inquiring after its very possibility, so it is highly misleading label Kant’s question or undertaking ‘semantic’.<sup>77</sup> While the premise of Kant’s transcendental investigation is semantic in character, the nature of the investigation itself, the question or problem which he is interested in, falls completely outside the purview of philosophical semantics. Indeed, arguably it is a problem which has no analogue in contemporary philosophical discourse at all (Stroud, 1977: 238).<sup>78</sup>

The final point to be made about Kant’s question concerns synthetic a priori judgements. Recall that at the end of the previous section, I explained that although Kant considers there to be a connection between the questions:

- (i) Under what conditions is experience possible a priori?
- (ii) How are synthetic a priori judgements possible?

---

<sup>77</sup> Arguably, insofar as Kant’s question concerns the very *possibility* of semantics, it would be more appropriate to label it ‘meta-semantic’.

<sup>78</sup> A clear exception to this is Hinzen (2009; 2012) and Hinzen and Sheehan (2013). The similarities between Kant’s transcendental undertaking and Hinzen’s project are developed fully in Chapter seven.

, elucidation of this connection is only possible once the relevant notion of experience is clear. Since this has now been clarified, the connection is, in turn, transparent. A priori conditions for the possibility of experience are the conditions of intentionality: conditions which must be met if any intentional relation to an object is to be possible. In this sense, all objects must conform to these conditions if they are to be cognised at all. Yet this seems to supply precisely what is required with respect to synthetic a priori judgements. Judging that these various conditions apply to objects is a synthetic a priori matter. It is a priori because these conditions apply necessarily, as opposed to merely contingently, to all objects. That is, it is not merely that objects just so happen to conform to these conditions, they must conform to these conditions if they are to be cognised at all. Furthermore, this is also a synthetic matter since these conditions do not reduce to the law of non-contradiction. Consequently, to undertake a transcendental “deduction” (A85/B117) of a priori conditions for the possibility of experience is also to undertake a deduction of the possibility of, or rather, the legitimacy of, a certain class of synthetic a priori judgements, thereby securing the possibility of metaphysics.<sup>79</sup>

#### 4.3. *The discursivity thesis*

One of Kant’s most fundamental claims about experience is that it is discursive: it requires *both* the sensible faculty and the faculty of the understanding. “Only from their union [sensibility and understanding] can cognition arise.” (B76) Kant characterises the faculty of the sensibility in terms of *receptivity* or *passivity*. Dickersen (2004: 6-7) offers the following metaphor for how this faculty should be understood. Imagine a globe of soft plastic. This globe can be acted upon by external forces which produce changes in the shape of the globe. We might then say that the globe has a capacity to be affected by its environment, with the shape changes constituting the matter of this capacity.<sup>80</sup> This is essentially how Kant understands sensibility. It is a faculty of receptivity insofar as it amounts to a capacity to be shaped or moulded by external objects. Kant repeatedly talks of sensibility yielding

---

<sup>79</sup> The price that has to be paid for securing this firm foundation for speculative philosophy, is that *all* of our knowledge (not simply that delivered by metaphysical inquiry) is knowledge of appearances: knowledge of objects as they conform to our mode of cognition. The mind-transcendent status of the objects of knowledge is sacrificed at the altar of metaphysical inquiry.

<sup>80</sup> Dickersen continues the metaphor with the idea that there is a viewer inside the globe which can only observe the internal changes in the surface of the globe. However, this strikes me as a needlessly tendentious extension: one which invokes the charge of homuncularism, as Dickersen himself notes. The crucial element of sensibility is that it is a faculty of determination; this is perfectly coherent without any additional ‘viewer’ inside the globe.

modifications (*Modificationen*) or determinations (*Bestimmungen*) of the subject (e.g. Corr: Letter to Beck, Jan. 1792, 400): always invoking the idea that these are no more than imprints left upon us, the purely passive recipients. These determinations are what Kant calls sensations: roughly, atomistic sense data of a particular phenomenal character such as taste, colour (A28), heat, sound (B44) and weight (A169/B211). A multitude or array of sensations Kant calls a manifold. In the *Transcendental Aesthetic* Kant argues that space and time are a priori forms of our sensible faculty. A stream of uncoordinated sensory data, or a manifold, is always grasped under the general forms of space and time. A manifold of sensations grasped under these forms is known as an *intuition*.

That Kant divides the cognitive labour between this sensible faculty and a separate faculty of understanding is, of course, perfectly consistent with the accounts of mental faculties offered by his predecessors. For example, we might think of it as mapping onto Hume's distinction between impressions and ideas. Or more generally, we might think of it as nothing more than a statement of the distinction between perception and thought. Such a benign assessment would, however, be deeply mistaken. In fact, the discursivity thesis actually imbues Kant's distinction of these two faculties with a deep significance. Discursivity entails that the sensibility and the understanding have *qualitatively* distinct and irreducible cognitive functions (A52/B74). If all cognition requires the contribution of the understanding over and above that of sensibility, the former faculty must contribute to or determine cognition in a way that sensibility alone is incapable of.

The contrast with Hume should be evident. For Hume, all cognition is constituted either by an array of sensible impressions, or by ideas which are simply dimmer copies of impressions. A sensory state is the *epitome* of a cognitive state for Hume, and the paradigm according to which all cognition is understood. In this sense, "Hume assimilated thinking to sensing." (Bennett, 1966: 54-5) Even setting Hume's sensory outlook however, the idea that the distinction between the intellectual and sensible faculties is merely one of degree, in particular, degree of epistemic clarity, was widely held by Kant's predecessors. For Leibniz, for instance, inversely to Hume's view, the function of the understanding was to exercise "merely a clarificatory function (bringing conceptual clarity and distinctness to what the senses present obscurely)" (Allison, 2004: 15). By arguing for discursivity, Kant was breaking with this heritage of viewing the intellectual and sensible faculties as each, taken in isolation, capable of yielding kinds of experience (Pippin, 1982: 37) (A271/B327). It entailed

that the sensible faculty, taken alone, is incapable of yielding any experience whatsoever (Pippin, 1982: 39).

What then is the faculty of the understanding for Kant? Why must it be invoked in any account of cognition? I think that it is productive to approach this question indirectly: namely, by showing why sensibility alone does *not* suffice for cognition. Approaching the understanding via this circuitous route allows us to, (a) undertake a closer inspection of the nature of intentional cognition, which allows us to, (b) clarify precisely why cognition cannot be purely sensible, which allows us to, (c) define the role of the understanding ‘by remainder’. (c) will however, be left for Chapter five. Consider Kant’s assertion that:

“...a manifold’s *combination* as such can never come to us through the senses; nor, therefore, can it already be a part of what is contained in the pure form of sensible intuition” (B129-130)

“...nothing composite can *as composite* be given to us – rather, *the composition* of the manifold is something we ourselves must always *produce*” (Corr: Letter to Beck, July 1792, 421)

This claim, that sensibility cannot compose or that composition cannot be simply received, is among Kant’s most famous. Despite this however, it is far from transparent what it means. I think that this composition claim can be illustrated by reference to a sensory array of the kind:

Figure 4.



Were sensibility capable of yielding intentional experience, then such experience could be thought of as essentially imagistic in this fashion.<sup>81</sup> That is, a relation to an object would be established by the imagistic nature of our experiential states or representations. This is essentially what Hume supposed.

There are insurmountable problems with this idea. One especially troublesome problem concerns individuation. If relation to an object is somehow established by the imagistic nature of the relevant states, then, necessarily, it must be possible to isomorphically map images to intentional states. That is, if intentional states are iconic, it must be possible to individuate images as finely as intentional states. Yet this is precisely what is not possible. Specifically, this possibility is foreclosed by a simple, yet powerful point: images are *ipso facto* non-intensional. For example, the president of the United States is the first black American president. Nevertheless, despite being co-extensional, attributing to someone a state with the intentional content ‘the president of the United States’ is distinct from attributing to them a state with the content ‘the first black American president’, i.e. they are intensional states. It follows from this that if images are to be individuated so as to correspond isomorphically to intentional states, an image of the president of the United States must be distinct from an image of the first black American president.<sup>82</sup> However, no such distinction is available. For example, consider the above image. Does it have the content ‘the first black American president’ or that of ‘the president of the United States’? The answer is either: neither or both.

This claim needs some unpacking. It is natural to suppose that determining whether the content of an image is that of ‘the president of the United States’ or of ‘the first black American president’ depends upon the intrinsic properties of the image in question. As such, it might be pointed out that while perhaps the above, rather impoverished image makes no such distinction, this does not preclude richer images doing so. However, to think this is to misunderstand the claim, which is, at its core, a principled one regarding the nature of iconic representation. Take *any* image, all the way from the impoverished one above to a far richer one which resembles the president of the United States a great deal, and ask: is it an image of the president of the United States or is it an image of the first black president of America? As should be clear, responses to this question do not turn at all upon the particulars of the

---

<sup>81</sup> This is clearly something of an oversimplification since not all sensations are visual sensations, and Kant certainly recognises this (B44). Nevertheless, visual sensations have a central role here, so I will use images as the central point of reference in this discussion.

<sup>82</sup> That is, there must be a way of distinguishing the intentional content of the images.

image(s) considered. Rather, for any image at all, whether it resembles the president of the United States a great deal or very little, *if* it resembles the president of the United States, then it also resembles the first black American president and vice versa. In turn, if an image does *not* resemble the president of the United States, then it does not resemble the first black American president either, and vice versa. What this shows is that iconic representation leaves no space for the representation of an object under *distinct* co-extensional characterisations. Images are not fine-grained enough to distinguish co-extensional descriptions (Fodor: 2007; Hinzen and Sheehan, 2013: 85; Landy, 2008). Yet, given that intentional states *are* this fine-grained, this forecloses individuating images so as to correspond isomorphically to intentional states.

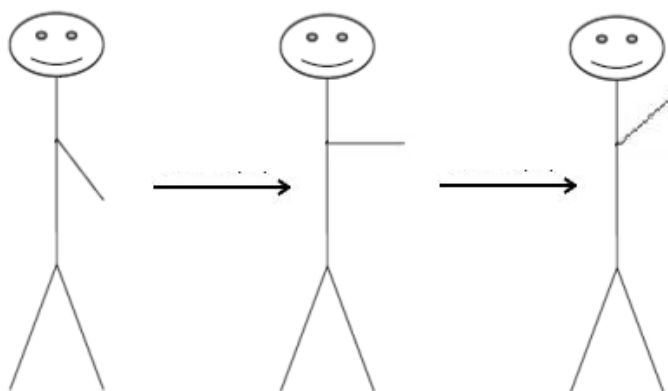
Consider another example. There is a difference in being in an intentional state about my wife and being in an intentional state about Judy's daughter, even where my wife is, as a matter of fact, Judy's daughter: these are intensional states. Now, suppose I own a portrait of her. Evidently, it would be nonsensical to insist that it is a portrait of my wife but not of Judy's daughter. The picture makes no such distinction. If it resembles my wife then it also resembles Judy's daughter. Yet clearly things do not stop here. If it is a picture of my wife then it is also a picture of a doctor who works at a Newcastle hospital, a picture of Emma's sister etc. In other words, images *badly* underdetermine intentional states. It is not merely that images cannot *\*quite\** be individuated as finely as intentional states: they are poles apart.

What this shows is that were relations to objects established by images, if the intentional content of our experiential states were iconic, there would be no intensionality. Our intentional states would be as fine-grained, or rather, as coarse-grained as the objects at which they are directed (Allison 2008: 111). Since however, this is *not* the case, sensibility alone cannot be capable of yielding intentional experience, which, for Kant at least, is fundamentally intensional.

It is worth noting that this problem is exacerbated when the relevant sensory array is extended over time, comprising successive states of the subject. Landy (2008) offers the example of a person raising his hand. Simplifying the perceptual manifold which would be involved somewhat, we could analyse this in terms of three successive sensory manifolds:



Figure 5.



Suppose that the man raising his hand is asking a question, such that ‘the man raising his hand’ and ‘the man asking a question’ are co-extensional. Despite this, clearly there is a difference between intentionally representing the man as raising his hand, or as asking a question (one could be in an epistemic position to grasp the former but not the latter). Given this, does the above sequence of manifolds represent the man as raising his hand or as asking a question? Evidently, the state of affairs here is no different from the previous iconic examples. Given that these descriptions are co-extensional, if the sequence resembles the man raising his hand then it also resembles the man asking a question. The image leaves no space for representing one, but not the other.

As it is sometimes said then, there is a difference therefore between a ‘mere’ sequence of representations (as above) and the representation of a sequence of representations *as a sequence*. Insofar as intentional states are intensional, temporally extended sequences of representations are represented *as* sequences in the sense that they are determined according to some characterisation or description e.g. a man raising his hand, or a man asking a question. Yet a mere temporal sequence of manifolds falls short of this. Again this shows that there is no prospect of individuating images, or in this case, series of images, as finely as the relevant intentional states.

To tie this back into Kant’s ‘composition talk’, it should be clear that this individuation problem can be understood as a problem of composition. For Hume, simple impressions and ideas fit together spatially and temporally to compose complex impressions and ideas respectively; they are individuated on this basis. However, what the foregoing observations show is that no such notion of spatial or temporal composition could ever suffice to individuate images finely enough to map them isomorphically onto intentional

states. As Kitcher says (1990: 79), spatio-temporal contiguity is too ‘promiscuous’. Consequently, the composition exhibited by experiential states cannot be given or received: it cannot be a function of the sensibility.

There is a separate argument against sensibility being sufficient for experience. Examples of sensations which Kant offers include, taste and colour (A28), heat and sound (B44) and weight (A169/B211). In other words, sensations or ‘presentations’ are sensory qualities: visual qualities, tactile qualities etc. A multitude or *manifold* of these qualities is always necessarily apprehended under the sensible forms of space and time. Now, crucially, Kant thinks that there is a logical gap between receiving sensations of this kind and relating them to objects or being in intentional states.

“We have within us presentations of which we can also become conscious. But no matter how far this consciousness may extend and how accurate and punctilious it may be, they still remain forever only presentations, i.e. inner determinations of our mind in this or that time relation. How is it, then, that we posit an object for these presentations; or how is it that in addition to the subjective reality that they have as modifications [of the mind], we also attribute to them who knows what sort of objective reality. Their objective significance cannot consist in the reference to another presentation... For otherwise the question returns: how does this other presentation, in turn, go beyond itself and acquire objective signification in addition to the subjective one that it possesses by being a determination of the mental state?” (A197/B242)

According to Kant, it is possible to separate or isolate the reception of sensations from their being related to an object. For instance, he speaks of sensibility yielding a ‘stream’ of subjective states (A107), comprised of arrays of sensations. The idea seems to be that the mere operation of sensibility can be paraphrased as ‘now, sensory qualities AA’, ‘now sensory qualities AB’, ‘now sensory qualities BB’ etc. That is, merely receiving sensory representations amounts to no more than a stream or inter-play of subjective states. No connection to an object is forged. Similarly, Travis (Ms.: 11) asserts that sensibility, taken in isolation, keeps us locked in our own inner world. In other words, it is possible to attribute a capacity for receiving sensory representations without this entailing the attribution of a

relation to an object.<sup>83</sup> This connection to an objective domain which exists independently of being represented is a *further* cognitive achievement.

Kant's claim can, I think, be formulated in what are perhaps more attractive terms. Sensations for Kant, recall, are modifications or determinations of the subject: no more than imprints left upon us, the purely passive recipients, by external objects. As such, another way of formulating his claim is that there is a logical gap between being *determined* in certain respects, and those determinations relating to objects. His claim is that the mere capacity for receptivity does not, in itself, account for the possibility of intentional experience. Representations being intensionally directed upon objects which exist independently of being so represented is not something which falls out simply from the capacity to be affected by our environment. Interestingly, this marks a repudiation of Kant's pre-critical position outlined in the *Inaugural Dissertation*. According to this earlier position, relation of representations to an object is explained by the subject being affected by the relevant object: that is, sensations represent an object insofar as they are the effects of objectual causes. This causal relation is said to explain how a modification of our mind can *represent* something or have a relation to an object (see, Longuenesse, 1998: 18-19). In repudiating this view, Kant is denying that intentionality can be accounted for in causal terms.<sup>84</sup>

Furthermore, Kant rightly believed that the difference between receiving sensory presentations and relating them to an object could not be explained by appeal to association.

“But suppose that I inquired more precisely into the reference of given cognitions... and that I distinguish it, as belonging to the understanding, from the relation in terms of the laws of the reproduction imagination (a relation that has only subjective validity).” (B141)

---

<sup>83</sup> Interestingly, this contrast shows up in Frege also. Frege distinguishes the ‘inner world’ of sense impressions - imagination, moods, inclinations - from *thought*. Sense impressions “alone do not disclose the outer world to us... To have a visual impression is not to see things... Having visual impressions is certainly necessary for seeing things but not sufficient. What must be added is non-sensible. And yet this is just what opens up the outer world for us” (1956: 308-309). Of course, Frege's non-sensible ingredient which unlocks the world for us - a thought from the ‘third realm’, is very different from Kant's response (Bell, 1979: 122). Moreover, there is good reason to think that appealing to non-psychological, ontological propositional entities in order to explain the objectivity of experience, as Frege does, is destined to fail (see King, 2009). Unfortunately, this issue takes me too far afield to justify proper consideration here.

<sup>84</sup> For this reason, I think Kant would be highly critical of the ‘causal theories’ of intentionality which are commonly espoused nowadays. *Crucially*, this is not to say that he considers causality to be irrelevant. Discursivity states that intentionality requires both the understanding and the sensibility, and since the latter yields, essentially, causal effects of objects, he absolutely has a place for causality in his account. Nevertheless, Kant's concern with intentionality only makes sense in the context of the claim that causality does not suffice for intentionality. As such, although he can grant the necessity of causality to intentionality, the notion of a ‘causal account’ of intentionality would be a misnomer to him.

Once the logical gap between a sensory stream and relating representations to objects is admitted, association does nothing to close it. Why would the fact that certain representations become associated explain how they acquire relation to an object? The law of association serves to move the mind from one kind of representation to another with which it has been constantly conjoined, for example, from the impression of weight to the representation of a body. It is tempting to think that in forging a connection between representations, association is forging an intentional state such as:

- Bodies are heavy

However, this is only plausible insofar as the respective representations are *already* assumed to bear a relation to an object. If certain representations relate to bodies, and certain representations relate to the objective property of weight, then perhaps it is reasonable to suppose that their association will amount to, if not at least be the grounds for, the objective state ‘bodies are heavy’. In this case however, association is not explaining the relation to an object, but presupposing it, and thus begging the relevant question. By contrast, if this relation to an object is not presupposed, there is no reason to think the association of the relevant representations could establish it. It would result only in a descriptive generalisation about my subjective state:

- {impression of weight, representation of a body}

That is, ‘whenever I feel the impression weight, I also have the representation of a body’. This evidently falls short of objective experience.

Although commentators have widely recognised that Kant is critical of the role allocated to association in the work of many of his predecessors (e.g. Fodor, 2003: 93), this criticism has seldom been related to the central issue of intentionality. In fact, occasionally this connection has been actually denied outright. For instance, Longuenesse (1998: 186-7) considers and promptly rejects understanding Kant’s critique of Humean association in this way. The reason, she says, that Kant’s critique of association is *not* that it cannot explain the possibility objective cognition over and above merely subjective cognition is that “of course, Hume does have a story to tell about how we come to believe that our impressions and ideas represent properties of the objects “whatever the state of the subject””. Longuenesse is right that Hume has a relevant associational story to tell. What Longuenesse omits to mention, or fails to recognise, is that this story is a comprehensive failure and tacitly presupposes what is

in need of explanation, as discussed in Section 2.2. In fact, far from showing that objective experience can be explained in terms of association, the failure of Hume's account seems to confirm the impossibility of such an explanation. Hume utilises association subtly and with considerable clarity of thought, and so the fact that he was unable to plausibly show how it might yield objective experience suggests that the shortcoming lies not in the detail of his account, but with the idea of association itself.

Kant's conviction is that objective cognition could not arise simply from a sensory array or, alternately, that a purely sensory array could never be objective. No matter how often sensory representations are given to the subject in specific patterns and configurations, this could not account for how they are taken to relate to an object. It is worth emphasising that it follows from this that sensibility is not only insufficient for veridical experience, it is also insufficient for illusionary experience:

“...truth *and illusion* are not in the object in so far as it is intuited” (A293/B350) [my emphasis]

In both veridical and non-veridical experience, representations or presentations ‘go beyond themselves’ (A197/B242) in purporting to represent an object other than themselves. Kant's claim is that sensibility alone cannot account for this.

“The basic contrast with which Kant operates in both the *Analytic* of the first *Critique* and the corresponding segment of the *Prolegomena* is between perception and experience. Merely having a series of perceptions is not equivalent to having experience.” (Allison, 2008: 206)

Consequently, experience is not possible merely through our receptivity to the world alone, but must necessarily depend upon the contribution another faculty also: the understanding. This insight entails a qualitative distinction of the functions of the sensibility and of the understanding, and is widely influential to this day:

“...to identify something as an impression is to place it in a logical space other than the one in which talk of knowledge – or, to keep the general case in view, talk of world-directedness, knowledgeable or not - belongs” (McDowell, 1994: xv)

“...the possibility of thought depends upon the idea of objective truth, of there being a way things are which is not up to us. I do not see how any causal story about the

sequence of stimuli reaching an isolated creature can account for this feature of conceptualisation or intensionality.” (Davidson, 2004: 141)

None of this is to deny the importance of sensibility. On the contrary, Kant insists that sensation is necessary for any experience:

“Now, there are two conditions under which alone the cognition of an object is possible, first, *intuition*, through which it is given, *though only as appearance*. (A92-93/B125)

The challenge is to explain how *by means* of given sensory data, we come to represent an objective order: an objective order that is not identical with the relevant sensory data (B234-235). Relating representations to objects is the characteristic activity of the understanding (Allison, 2004: 173). Again borrowing Travis’ formulation (Ms.: 11), it is the contribution of the understanding which serves to ‘unlock’ the objective world for us.

#### **4.4. Global experience**

Perhaps because the most profound and revelatory consequence of the discursivity thesis is that sensible representation alone is insufficient for experience, there has been a prevailing tendency to portray Kant’s topic as being primarily, if not entirely, that of *perceptual* cognition (e.g. Bird, 1962: 82). After all, *prima facie*, it certainly seems plausible that one can be in a perceptual state, one which is marked by intentional content and therefore amounting to experience, without any involvement of a non-sensible faculty. As a consequence, since discursivity denies this *prima facie* attractive thesis, the focus has tended to reside primarily, if not exclusively, with perceptual experience. Kant’s challenge is thus commonly understood solely in terms of determining what perception presupposes over and above what the sensibility can provide. Moreover, this emphasis upon perception has been reinforced by the way in which ‘experience’ is used in ordinary language. Kant’s transcendental investigation is said to concern the a priori conditions for the possibility of *experience*. Yet, ‘experience’ is

a term generally used to denote perceptual apprehension of the world. For instance, ‘having an experience’ might refer to jumping out of a plane or tasting an exotic food.<sup>85</sup>

*However*, given that ‘experience’, for Kant, denotes any intentional mental state, the identification or restriction of experience to perception does not appear viable. After all, perceptual states are clearly not the only kind of mental state which exhibit intentionality. Contemporary philosophers generally recognise a wide range of intentional states or, so-called, ‘propositional attitudes’: thinking, believing hoping, desiring, etc. (Searle, 1983). Relative to perception, what unites all these postures of the mind is that (a) they are intentional and (b), they are not perceptual; they do not directly involve any sensory manifold. As such, they represent an object *in absentia*.

It is clear that Kant does take experience to extend beyond perceptual states. For instance, he speaks of cognition of the inhabitants of the moon:

“That there could be inhabitants of the moon, even though no human being has ever perceived them, must of course be admitted.” (A493/B521)

Given that in the eighteenth century there was little prospect of perceiving inhabitants of the moon, the fact that Kant finds it possible to think of or speak *about* such inhabitants, shows that objective cognition is not limited to perception. Similarly, Kant also speaks of events which occurred at a time prior to one’s own existence (A495/B523), which obviously raises the same issue. In other words, Kant accepts that there is such a thing as non-perceptual experience (Kukla, 2011; Pendelbury, 1995: 779; Pereboom, 1988: 340-1; see, Melnick, 1993, for a critique of Kitcher’s neglect of this point). Nevertheless, Kant makes no attempt in CPR to sub-divide non-perceptual experience into anything like the rich classifications of propositional attitudes so prominent in contemporary philosophy, so I shall speak only of non-perceptual experience generally. Following, Kukla (2011), non-perceptual experience will be called ‘global experience’. Borrowing Kant’s terminology then (A320/B376-377), experience in general is the genus, under which stands both perceptual experience and global experience.<sup>86</sup> Consequently, insofar as the transcendental method is concerned with the a

---

<sup>85</sup> Though ‘cognition’, which commentators use interchangeably with ‘experience’, lacks these strong sensible or perceptual connotations.

<sup>86</sup> This division is, I think, what Kant has in mind when he says that: “...it is empirical cognition if the object is given in the senses’ representation (the latter includes both sensation and sensation bound up with consciousness i.e. perception); it is a priori cognition if the object is given, but not given in a representation of the senses.” (Corr: Letter to Beck, Jan. 1792, 400)

priori conditions of experience, this encompasses the conditions for both perceptual and global experience.

It might be objected that global experience is in tension with Kant's discursivity thesis. Specifically, discursivity entails not only that all experience involves the understanding, but also that it requires the sensibility. Yet, defining global experience in contradistinction to perception seems to amount to a denial that there is any sensible element inherent to the former. However, this tension can be dispelled by consideration of some of Kant's formulations of discursivity:

"... all thought, whether straightaway or through a detour, must ultimately be related to intuitions, thus, in our case, to sensations, since there is no other way in which objects can be given to us" (A19/B33)

"The postulate for cognizing the actuality of things requires perception... - not immediate perception of the object itself the existence of which is to be cognized, but still its connection with some actual perception." (A225/B272)

In these formulations, discursivity is stated so as to be consistent with the possibility of non-perceptual experience. Kant posits that experience can relate directly to sensations or relate to sensations *indirectly*: either route is said to satisfy discursivity. Yet, this seems to be exactly what is required for global experience. Specifically, global experience can be distinguished from perception by not relating to sensations directly, but can nevertheless satisfy discursivity insofar as it relates indirectly to sensations. Of course, at this point we do not know what this 'indirect' relation amounts to, but the mere fact that Kant postulates such a relation is indicative of a distinction between perceptual and non-perceptual experience.

Finally, it is worth guarding against a possible misapprehension. Just because Kant grants that there are non-perceptual, experiential states, this clearly does not entail that all mental states are intentional or constitute experience. This is certainly not Kant's view (Longuenesse, 1998: 66; see Section 5.3 for discussion). Rather, he is simply committed to the view that the states which *are* intentional or experiential extend beyond those of perception.



## Chapter 5. Rules and Normativity

---

The preceding chapter concerned the nature of experience: how it is understood in terms of intentionality, its role within Kant broader transcendental philosophy, and why it is not reducible to sensibility but requires also the contribution of the understanding. However, beyond gesturing towards this intellectual faculty, nothing was said of how Kant seeks to account for or explain intentional experience in a positive fashion. This account is the central topic of this chapter; it can be broken down into three tightly connected claims:

1. Experience is essentially rule-subjecting.
2. Rules presuppose a capacity for judgement.
3. Judgements exhibit a finite number of forms.

The first two claims are very intimately related, so much so that while (1) is introduced in Section 5.1, it can only be fully developed and explicated in the context of (2): the topic of Section 5.2. Section 5.3 clarifies the position by addressing some potential problems and complications. Finally, Section 5.4 is concerned with (3); specifically, how we should understand, or what the significance is, of the forms exhibited by judgements.

### *5.1. The normativity of the intentional*

“...our thought of the relation of all cognition to its object carries something of necessity with it, since namely the latter is regarded as that which is opposed to our cognitions being determined at pleasure or arbitrarily... insofar as they relate to an object our cognitions also necessarily agree with each other in relation to it, i.e. they must have that unity in which the concept of an object consists.” (A104-5)

How are we to understand Kant’s claim that experience “carries something of necessity with it”? Evidently, this necessity cannot be equivalent to the epistemological necessity of a priority, for this would entail that no experience could be justified a posteriori: something Kant certainly does not accept. On the contrary, far from being dichotomous with the a posteriori, Kant stresses that this type of necessity is exhibited by empirical, and therefore, contingent, experience (B142) (cf. Bird, 1962: 116-7 & 143). Rather than being an

epistemological notion, Kant is referring to a different type of necessity, that of being somehow “non-arbitrary”. All intentional experience is said to exhibit this latter type of necessity.

In order to try and explicate what Kant has in mind here, consider the distinction between a purely descriptive law and a normative law. The latter describes how things should be, whereas the former describes certain regularities in the order of things. The difference resides in the opposing directions of fit between the laws and the facts they range over. A descriptive law is valid only insofar as the facts it ranges over conform to it. If there is a violation of the law, we adjust it to accommodate the facts. In contrast, a normative law remains binding even when it is violated. For example, the force of ‘you ought to prepare for your philosophy tutorials’ is not weakened or undermined by the persistent flouting of it. Accordingly, normative laws demand or *necessitate* conformity from that over which they range. It is this idea of being subject to a normative law or a *rule* which Kant means by cognition carrying ‘something of necessity’ (see, e.g., Allison, 2004: Chpt. 7, fn. 42).

“Suppose that we inquire what new character is given to our presentations by the *reference to an object*, and what is the dignity that they thereby obtain. We then find that this reference does nothing beyond making necessary the presentations being combined in a certain way and being subjected to a rule.” (A197/B242-3)

“Even if the rule does not leap to the eye, one must still represent the object as in accordance with a rule in order to conceive that it represents something, that, is which has a certain place and function among its other determinations.” (Kant, *Reflexionen*: in Guyer, 1987: 74)

One of Kant’s master ideas is that for representations to relate to an object or for experience to be intentional, is for it to be rule-subjecting: that is, it subjects the *experiencer* to rules.<sup>87</sup> This is what representational purport amounts to for Kant. It is to attribute to experience the normative ‘necessity’ of subjecting one to rules. Given this, and given that, as was noted earlier, Kant defines the role of the understanding as that of bringing representations into relation with an object, it is no surprise that the understanding is defined as the faculty of rules (A133/B172 & B197-8). Further, since Kant also speaks of the understanding as the faculty of concepts, it follows that concepts simply *are* rules:

---

<sup>87</sup> I shall speak of experience being ‘rule-subjecting’. This should be taken as shorthand for experience subjecting *the subject* to rules.

“All cognition demands a concept, though that concept may, indeed, be quite imperfect or obscure. But a concept is always, as regards its form, something universal which serves as a rule.” (A106)

As such, the following equivalences hold for Kant:

Faculty of the understanding = Faculty of concepts = Faculty of rules

Unfortunately, despite its importance, Kant offers little detailed discussion of how experience is rule-subjecting.<sup>88</sup> Most of this section is devoted to developing this idea with some precision, with reference to examples (a tool Kant himself employs with notorious infrequency). Having a clearer picture of what the rule-subjectingness of experience amounts to in hand, I shall return to Kant’s master idea that the intentional is essentially normative at the end of the section.

Consider first, rule-subjectingness with respect to perceptual experience. The starting point for perceptual intentionality, as discussed in Section 4.3, is a succession of inner determinations of our mind, manifolds of sensibility, in this or that time relation. The challenge is to say how in addition to their reality as subjective modifications, they come to have objective signification. There is a significant distinction between two different kinds of perceptual experience which needs to be drawn from the outset. Kant points out that while a manifold, as a modification of the subject, is always successive, the objective signification of a manifold is either successive or sequential (A182/B225). For example, take the perception of a house ‘standing before me’. In such a case, it is likely, or at least possible, that the relevant sensations be apprehended sequentially. The order of my apprehension could run from the roof of the house to the bottom, or from the left of the house to the right, or even be spread over a considerable time as, say, I walk slowly around the house surveying it. The crucial point is that many such sequences of sensations are possible: there is no particular successive order to which they must adhere for the perceptual experience *to be that of a house*. The order in which the relevant sensations are apprehended is inessential to the object of perception itself. By contrast, consider the perception of a boat floating down river. Just as in the house example, there is a succession of sensations: boat upstream-boat downstream. Yet, unlike the house case, the particular successive order of these sensations is essential to

---

<sup>88</sup> The limited discussion of the matter in CPR itself has received little elaboration in the secondary literature. As Landy (2009: 27) notes, the question of what Kant really means by ‘rules’, and by the claim that experience is rule-subjecting, has generally not been a major topic of interest for commentators.

the object of perception. If what was perceived was instead: ship downstream-ship upstream, one would be perceiving a different object altogether: namely a ship floating upstream. In other words, by this successive-sequential distinction, Kant means to underline the difference between the perception of an event and non-eventive perception. This presents a puzzle because in both cases, what is apprehended is merely a sequence of sensations, A-B: so, this cannot account for how in event perception this sequence is taken to be determined in the object as successive i.e. an event. Something else must secure this difference. It is important not to conflate these kinds of cases because Kant ultimately wants to say that there is a difference in how they are rule-subjecting which accounts for the difference. Nevertheless, they are similar enough that I shall discuss the rule-subjectingness of perception primarily with reference to sequential or non-eventive perception, subsequently amending, or rather, extending, the account as required for successive or eventive perception.<sup>89</sup>

In the perception of a house, says Kant, the manifold is synthesised according to a concept or rule: that of ‘house’. Our sensory impressions are brought under this rule. According to Kukla (2011: 134), we should understand this rule in terms of what can be encountered in ‘investigative behaviour’, and I think there is something correct about this. Any perceptual experience, whether it be temporally extended or only a time slice, will be from a certain perspective or vantage point, in the sense that it implies an array of investigative, behavioural possibilities for modifying the experience. For example, if I perceive the front of a house, I can walk around to the back of the house; if I perceive the screen of my laptop, I can close my laptop; if I perceive a cup of tea on a table, I can taste the tea. In each case, the perceptual experience implies behavioural or investigative possibilities for changing the experiential state.<sup>90</sup> It seems reasonable to suppose that any perceptual experience will imply investigative possibilities of this sort. Of course, there may be physical impediments which constrain their realisation. However, what is relevant here is not the realisation of such investigative possibilities, that the behaviour is actually performed, but rather, *beliefs regarding* these possibilities. For example, if I perceive the front of a house, I may believe that were I to re-position myself in a certain way, I would perceive the back of the house; if I perceive the screen of my laptop, I may believe that were I to close my laptop,

<sup>89</sup> Bird (1962) contains a very lucid discussion of this matter (see pp. 154-167).

<sup>90</sup> Describing this behaviour as ‘investigative’ is potentially misleading insofar as it suggests that the behaviour must concern a particular object or set of objects which are *investigated*: the tea is investigated by tasting it, the house is investigated by circling it. Although such examples best illustrate the notion of investigative behaviour, such behaviour need not be directed upon a particular object in this way. If I am enjoying a panoramic view, an investigative possibility could be simply how the view would change if I changed my position. Here, there is no object which is ‘investigated’, but I would still describe this as an investigative possibility.

I would perceive only the outer shell; if I perceive a cup of tea I may believe that were I to raise it to my lips, I would taste tea, rather than water or bleach. Notice that these are all conditional beliefs, in which the performance of certain investigative behaviour is specified as the antecedent.<sup>91</sup> The beliefs are about what would ensue under the condition of certain investigative behaviour. Call these investigative beliefs.

Investigative beliefs are relevant to how perceptual experience is rule-subjecting because part of what it is for perception to be rule-subjecting makes essential reference to such beliefs (Landy, Ms.: 13; Bennet, 1966: 126). Specifically, perceptual experience is rule-subjecting insofar as to be in a perceptual state is, *inter alia*, to be *committed* to certain investigative beliefs. For example, to have a perceptual experience of a cup of tea is, amongst other things, to be committed to, for instance, the investigative belief that if the tea is sipped, what is tasted will be tea. A perceptual experience of the screen of a laptop commits one to the investigative belief that if the laptop is closed, only the outer shell will be perceived.<sup>92</sup> The idea here is that it is not simply an incidental effect of perceptual experiences that they commit the subject to investigative beliefs in this way, but that this is essential to the perceptual experience being *experiential at all*. It is part of what marks the difference between a mere manifold and perception. A mere interplay of sensations does not commit the subject to any investigative beliefs. This commitment only arises once concepts are applied, and this is, in part, *why concepts are rules*. Every concept is a perceptual rule insofar as its application to a sensory array commits the subject to certain investigative beliefs. Of course, for any given concept, the range of investigative beliefs it may commit one to is very broad and possibly not finite,<sup>93</sup> and almost certainly going to be a rather uncertain matter; attempting to specify or enumerate this range even for a single concept is probably a hopeless task. However, this task is tangential to the present concern which is simply the claim that subjecting a manifold to a concept will necessarily involve a commitment to *some set* of

---

<sup>91</sup> I am speaking of investigative *beliefs*, yet it is plausible that perceptual experience implies other attitudinal states besides beliefs e.g. expectations. It seems to me that beliefs are primary – but this should not be taken to prejudice the question of perceptual experience implying other attitudinal states besides.

<sup>92</sup> Notice that it is the subject of a perceptual experience which is committed in these ways. It is, evidently, nonsensical to speak of the experience being committed or the perception being committed; it is the bearer or subject of the experience which is committed. This is far from trivial because it entails that if perception is essentially rule-subjecting in this fashion, then perception essentially makes reference to a subject capable of being committed in this way. This connection between rules and a subject of experience is actually the key node in Kant's account of apperception: rules and self-consciousness are interdependent and arise together: neither is prior to the other.

<sup>93</sup> Consider, for instance, how investigative beliefs can be stated disjunctively i.e. where the consequent is disjunctive.

investigative beliefs. To say that perceptual experiences essentially involve commitments to certain investigative beliefs is part of what it means to say that perception is *rule-subjecting*.<sup>94</sup>

Some clarifications of this thesis are in order. Firstly, as has already been stressed, the claim is not that the perception of a house, or indeed, any perceptual experience, demands or requires investigative behaviour: that I am, for example, required to verify that my experience is of a house by inspecting it from different angles. Rather, the claim concerns investigative beliefs: what my experience would be *were* I to undertake certain investigative behaviour. Secondly, while the claim does pertain to beliefs, it does not concern the course of one's psychological states or attempt to describe the path they will follow according to some psychological model. The claim is not that, for example, whenever I have a perceptual experience of a cup of tea, I will subsequently think, 'if I drink the tea, I will taste tea'. This would be to draw a descriptive generalisation, one which does not require the notion of a rule at all, but, perhaps, only the postulation of a Humean law of association operating on the constant conjunction of past impressions of the sight of tea and the taste of tea (see Longuenesse, 1998: 49-50). Instead, the claim is that if I perceive a cup of tea then I *ought* to accept that if I drink the tea, then I will taste tea. This is quite different from claiming that such a conditional will always be thought subsequent to any perceptual experience of tea. The claim is a normative one, not a descriptive one. Thirdly, while the claim does concern what ought to be believed, this does not pertain to occurrent beliefs. To say that a perceptual experience, E, commits one to certain investigative beliefs,  $B_1 \dots, B_n$ , is *not* to say that whoever is in state E, should think  $B_1 \dots, B_n$ , but instead that it is incorrect to be in state E, and to *deny* any of the investigative beliefs  $B_1 \dots, B_n$ . It is in this sense that experience is rule-subjecting. It *subjects* the subject of experience to rules regarding what ought to be believed and what ought not to be believed. To say that someone is committed in certain ways is to say that they are subject to rules.

So, a concept is a rule for perceptual experience insofar as its application to a manifold essentially commits the subject to investigative beliefs. This is not all however: the normative reach of concepts extends beyond investigative beliefs, to what I will call 'non-investigative' or 'direct' beliefs. Investigative beliefs are always conditional: if certain investigative behaviour, then P. Nevertheless, perceptual experience also involves

---

<sup>94</sup> While I am arguing that perceptual experience *commits* the subject to investigative beliefs, this three-place relation could plausibly be cast in different (though always normative) terminology. Instead of commitment, it is possible to interpret this relation in terms of prohibition or permissibility (Landy, 2009: 18). Nevertheless, univocity has its advantages, and I will speak only of commitment.

commitments to beliefs which are not conditional in this fashion: not dependent upon a condition of investigative behaviour. For example, to perceive a house is to be committed to certain existential beliefs about houses: namely, that there is at least one such object.

Similarly:

“... in the case of the perception of something outside of us the concept of body makes necessary the representation of extension, and with it that of impenetrability, of shape etc.” (A106)

The concept of body ‘makes necessary’ these representations not in the sense of requiring that one think *impenetrability*, *shape*, etc., but in the sense of committing one to certain beliefs regarding them, and prohibiting others. For example, to perceive a body is to be committed to the belief that what is perceived is impenetrable.

Notice that the two examples provided of normative relations to direct beliefs are, for Kant at least, analytic and so underpinned by the rules of logic. In the house instance, the move can be understood as a straightforward existential generalisation. The body example requires some conceptual decomposition, but granted that, the relation can be stated in similarly tautologous terms. Clearly however, not all normative relations to direct or non-investigative beliefs need to be of this sort. To perceive a house may, for example, commit one to the belief that someone lives in it, or to perceive a cup of tea may commit one to the belief that someone brewed it, yet there is no sense in which these relations are undergirded by logic. Following Sellars (1953), call this the difference between formal and material commitments. This is a vital distinction because it marks a difference of normative force. To say that a perceptual experience, E, formally commits one to a belief, B, is to say that E supplies one with an all-things-considered reason,  $R_a$ , to accept B: that is, there is no mitigating or extenuating circumstance which could render  $R_a$  an insufficient reason to accept B. If I perceive a body, then irrespective of whatever else may be the case, I have a sufficient reason to accept that there exists at least one thing which is a body: to have a reason of this sort is what it means to be *formally* committed. On the other hand, many commitments – material commitments – do not supply reasons of this sort. To say that a perceptual experience, E, materially commits one to a belief, B, is to say that E supplies one with a *pro tanto* reason,  $R_p$ , to accept B, in the sense that there *are* mitigating or extenuating circumstances which could render  $R_p$  an insufficient reason to accept B. For example, to perceive a cup of tea is to be, *inter alia*, committed to the belief that someone brewed it.

Nevertheless, it is conceivable, albeit wildly improbable, that no one brewed it, but that the tea fell from the sky into the cup. In these circumstances, i.e. where one believed that the tea fell from the sky into the cup, the reason to accept that someone brewed the tea supplied by perceptual experience of the tea, would not be a *sufficient* reason to accept that someone brewed the tea; as such, it is only be a *pro tanto* reason: to have a reason of this sort is what it is to be *materially* committed.

Since this ‘committed how’ distinction between formal and material commitments cuts across the ‘committed to’ distinction between investigative beliefs and non-investigative beliefs, it is possible to tabulate the rule-subjectingness of sequential perception as follows:

Table 1.

Committed-how	Committed-to
Formally committed (all-things-considered reason)	Investigative beliefs
Material commitment ( <i>pro tanto</i> reason)	Direct or non-investigative beliefs

It is an interesting question whether there is such a thing as a formal commitment to an investigative belief (all of the examples, I discussed earlier are, I think, examples of material commitments). Possibly there are not, but I set this issue aside.

So far, I have discussed only sequential perception. The account provided is largely applicable to successive perception also, for, like sequential perception, successive perception essentially involves both formal and material commitments to investigative and direct beliefs. Nevertheless, this account needs to be augmented in the case of successive perception. Why is it that in successive perception, the order of sensations is taken as determined in the object, whereas in sequential perception it is not? According to Kant, the difference resides in the nature of the rules applied in the respective cases:

“...in the perception of an event, there is always a rule that makes the order in which the perceptions follow upon one another (in the apprehension of *this* appearance) a necessary order.” (B238)

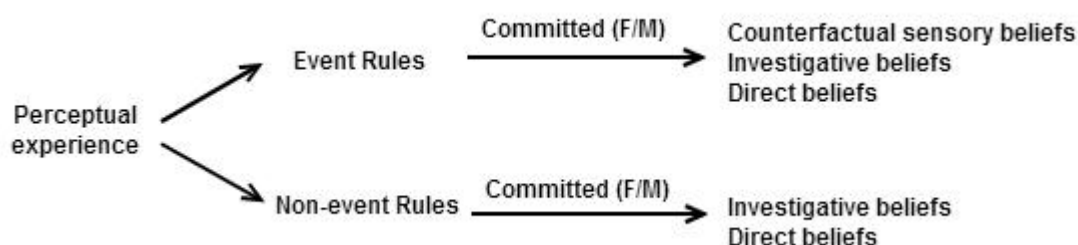
In other words, subsuming a manifold under certain rules, “compels us to observe this order of perceptions rather than some other order” (A196/B242), and it is in the sense that the order is taken to be determined in the object itself (A192/B237-8). The idea here is fairly straightforward. Basically, certain concepts are ‘event concepts’ insofar as their application to



a manifold invokes special commitments, absent in the case of ‘non-event concepts’. These commitments concern counterfactual beliefs of the form: ‘if the order of sensations had been different, the object perceived would have been different’. For example, to apply the concept ‘float downstream’ to a manifold is, in part, to be committed to the counterfactual belief that had the order of my sensations been other than what they were e.g. boat downstream-boat upstream instead of boat upstream-boat downstream, the object perceived would have been different also i.e. a boat floating upstream. Here, the antecedent of the counterfactual belief makes essential reference to the order of my sensations: this is the condition for the perceptual experience being of what it is, so call these beliefs ‘counterfactual sensory beliefs’. By contrast, to bring a manifold under the concept ‘house’, while essentially involving a range of commitments, none of these concern counterfactual sensory beliefs which refer to the *order* of my sensations. As such, it is not an event concept.

Given this, all concepts can be divided into event concepts and non-event concepts. For example, some instances of the former include: ‘float downstream’, ‘hit’, ‘fall’, ‘raise’ ‘lit’ etc. Some instances of the latter include: ‘house’, ‘car’, ‘green’ etc. With this division in hand, the account of perceptual experience described might be depicted as follows:

Figure 6.



From the perspective of rule-subjectingness, the move from perceptual experience to global experience is a non-trivial one. Specifically, whereas perceptual experience is marked by the presence of a sensory manifold, global experience is marked by its absence. To think that there might be inhabitants of the moon is perfectly intentional or experiential, yet there is no sense in which this could involve subsuming a sensory manifold under rules! Still though, insofar as global experience is *experiential*, i.e. objective, intentional etc., it must involve concepts, rules. Similarly, discursivity requires that, despite the absence of a sensory manifold, sensibility must also play a role in such experience. Kukla (2011: Chpts. 4 & 5) offers an articulated account of how global experience is rule-subjecting and although I think

that it is flawed, it nevertheless contains an important insight into how the matter ought to be understood.

According to Kukla, the representation of spatially or temporally remote state of affairs, i.e. global experience, involves the claim that it is *encounterable*.<sup>95</sup> Specifically, “cognition of unperceived objects requires representation of a possible sequence of experiences connecting one’s current perceptual states to the remote object.” (Kukla, 2011: 150) Experience of unperceived objects is possible insofar as one can represent a series or sequence of possible perceptions, beginning from one’s present perception, which would terminate in the immediate perception of the relevant object. In this way, global experience refers to a possible intuition, thereby upholding discursivity. For example,

“...my representation of the Nishitama district in Tokyo is the thought of a possible procedure (or experiential route) by which I can advance to that district from my current perceptions” (Kukla, 2011: 151)

Rules enter into this picture insofar as, for Kukla, rules for global experience *just are* procedures stipulating series of perceptions of this type. That is, concepts or rules are procedures for ‘getting somewhere’ or ‘getting to something’ from one’s current perceptual position. They are necessary for global experience insofar as cognition of a remote object presupposes precisely this kind of procedure. Simply put, I can think about something which is not here insofar as I have a concept which tells me how I could advance to the immediate perception of it i.e. a concept which links me to a possible perceptual state of the of the right kind. This ensures the encounterability of the object.

The obvious problem with this view is that aside from in a very small number of cases, no such procedure is available. For example, I can represent a procedure for returning home from my present state: follow the road for ten minutes and then turn right. However, in general, even for quite familiar locations, I often cannot represent such a procedure. Worse, if the account is problematic for familiar locations, it seems to foreclose the objective cognition of objects never previously encountered altogether. I certainly cannot represent a procedure for advancing from my present position to the immediate perception of the tallest man in the world. Yet, since global experience of such a remote object is precisely the kind of actuality

---

<sup>95</sup> For the purposes of simplicity I will discuss only spatially, not temporally remote objects here, though Kukla’s account purports to explain both possibilities. The reason for this is that, as I shall argue, Kukla’s account fails even for spatially remote objects, and this failure can be generalised without difficulty to temporally remote objects also.

Kukla is seeking to explain, it cannot be that it presupposes such a procedure. This seems to demand far too much.

Kukla attempts to evade this difficulty by adopting an extremely loose notion of a ‘possible procedure’. Rather than constituting a finite list of concrete steps for advancing to a perceptual state of the right kind, a possible procedure is far less onerous:

“Even if a distant corner of the universe lies outside the reach of any technological device, including the most advanced space telescopes, we can still think of a possible procedure by which to be brought into its perceptual vicinity.” (Kukla, 2011: 151)

Evidently, no one is able to represent a series of concrete steps for advancing to distant regions of space, especially when it is stipulated that the regions lie outside the reach of any technology. But then, in what sense can we, nevertheless, represent a ‘possible procedure’ for being brought into its perceptual vicinity? Well, certainly it is possible to *conceive* of being so brought. It is conceivable that a suitable space ship could be built or a sufficiently powerful telescope could be designed which would serve this purpose. At the limit, this is all Kukla seems to mean by the notion of a ‘possible procedure’ and as such, it is surely correct that they are available for any global experience. All global experience seems to entail the *conceivability* of being brought to perceptual contact with the relevant object. However, the problem with this is that we have completely lost sight of the sense of which global experience is *rule-subjecting*. On Kukla’s view, possible procedures amount to quasi-imaginative capacities: capacities to conceive of the possibility of being brought into immediate perceptual relations with objects. Yet there is nothing normative or ‘ruleish’ about this. Despite equating concepts or rules with such possible procedures it is clear that this cannot account for the sense in which global experience is rule-subjecting.<sup>96</sup>

Although Kukla’s account is flawed, I think that it nevertheless contains an important insight: that objects being *encounterable* is central to global experience. This is because it secures the connection to sensory manifolds or, at least, possible sensory manifolds which seems to be required by Kant’s discursivity thesis. The problem is that Kukla combines this idea with the view that concepts are procedures for advancing to remote objects. This

---

<sup>96</sup> Indeed, it might be pointed out that even the more robust notion of a ‘possible procedure’, a concrete list of steps for advancing to an intuition of an object, is not a rule in the normative sense of the term, or, at least, it need not be. There is nothing inherently normative about representation of a possible sequence of experiences connecting one’s current perceptual states to the remote object. In my view, despite his emphasis on concepts being ‘rules’ for Kant, the root of the shortcomings of Kukla’s account is that he fails to properly grasp the sense in which Kantian experience is normative. I discuss this issue further at the end of Section 5.3.

bleaches concepts of their normative core and so fails to explain how global experience is rule-subjecting. By contrast, my view is that non-perceptual experience is rule-subjecting insofar as it commits us to certain beliefs. That is, despite lacking a sensory manifold, the sense in which global experience is rule-subjecting can be understood in fundamentally the same terms as the rule-subjectingness of perceptual experience: all global experience involves concepts or rules which function so as to commit the subject to certain beliefs. The most obvious kind of belief global experiences commit one to is what I have labelled ‘direct’. For example, fairly trivially, any global experience concerning the inhabitants of the moon commits one to the belief that there are inhabitants of the moon. Just as with perceptual experience, commitments to such beliefs can be either material or formal. I do not think that non-perceptual experience involves commitments to either investigative beliefs or counterfactual sensory beliefs. It is difficult to see how they could. Nevertheless, it does involve commitments to a special class of direct beliefs. Specifically, global experience of an object essentially commits the subject to a belief of the form:

‘There is some spatio-temporal point such that, were it occupied, one would immediately perceive the relevant object’<sup>97</sup>

Call such beliefs ‘direct sensory beliefs’. In committing subjects to beliefs of this form, global experience essentially involves reference to a possible intuition. It commits one to the possibility of a sensory manifold of the object of the experience, i.e. there is some spatio-temporal point which is such that were it occupied, one would immediately perceive the relevant object. As such, commitments to these beliefs satisfy the requirement that global experience is encounterable. Importantly, in contrast to the idea of a ‘possible procedure’, there is no requirement that subjects be capable of representing a procedure for advancing to the relevant intuition from their present perceptual state. Rather, subjects of global experience must simply be committed to the possibility of such a sensory state.

With this detailed discussion of the kinds of rules experience subjects one to, we are in a better position to understand what I earlier called one of Kant’s ‘master’ ideas. This is that apprehending rules and moving in a normative space is not something built upon an independently given, intentional experience of the world. Rather, experience essentially

---

<sup>97</sup> I am classing these beliefs under the heading of ‘direct’ rather than ‘investigative’ or ‘counterfactual sensory’ for the obvious reason that they are categorical in form rather than conditional.

presupposes such a normative capacity from the outset.<sup>98</sup> It requires rules of the kind I have outlined.

“For Kant, the aboutness characteristic of representing is a normative achievement.”  
(Brandom, 2002: 23)

Kant’s master idea is that if one ‘stripped away’ the normative element of experience, its rule-subjectingness, the sense in which it commits the subject to certain things, one would be stripping away its relation to an object (see Allison, 2004: 181; Bird, 1962: 80). What would remain would be manifolds of sensation, streams of phenomenal consciousness in one configuration or another. Of course this ‘remnant capacity’ for manifolds could still constitute a significant form of intelligence; it may still enable one to navigate one’s environment ably.<sup>99</sup> Yet, it would not amount to *experience*. It is only with the addition of rules that we are able to ‘get outside of our own heads’ in the sense in which Kant both assumes is possible, and is interested in accounting for transcendently.

“...man is a creature not of *habits*, but of *rules*. When God created Adam, he whispered in his ear... “When you cease to recognize rules, you will walk on four feet.”” (Sellars, 1967: 298)

Crucially, as I have described the position, nothing has turned upon any rules *in particular*. Although I outlined the kinds of rules applicable to perceptual and non-perceptual experience respectively, the specific rules discussed, e.g. ‘house’, ‘floating downstream’, are evidently not of any special significance. As should be clear, the thought has been that experience must subject one to *some* rules, but not any *particular* rule, involve *some* commitments, but not any *particular* commitment. In this sense, the idea has been that experience must simply operate within a normative space.<sup>100</sup>

---

<sup>98</sup> There are contemporary philosophers who defend positions similar, at least *prima facie*, to the one I am ascribing to Kant - that intentional states are essentially normative. Indeed, Wedgwood (2009) is entitled ‘The normativity of the intentional’. Also, Nick Zangwill (1998; 2010) contends that at least some intentional states are essentially normative. However, the difficulty with connecting or comparing Kant’s position to such contemporary theses in any robust sense is that the latter are founded upon a crucial theoretical distinction unrecognised and, arguably, inimical to Kant: the distinction between act and content or what Sellars’ called the “ing/ed” distinction. For example, Zangwill’s thesis (1998) is that propositional attitudes, as opposed to the contentful objects of these attitudes, are essentially normative. Insofar as Kant fails to recognise the distinction of act and content altogether (cf. Rodl, 2009), it is difficult to compare his view to contemporary analogues.

<sup>99</sup> See Section 5.3

<sup>100</sup> In the end, Kant argues that there are a special, circumscribed class of rules which are necessarily applicable to all experience (see Sections 5.4 - 6.4). However, this comes later in the dialectic, and for now what matters is a more basic and minimal claim that experience is always rule-subjecting in one way or another.

The claim that operating within this normative space is integral to the objectivity of cognition strikes me as compelling. It seems plausible that did experience not commit the subject in certain ways, relation to an object would be lost. We would be stranded with our own mental presentations: bereft of any conception that our own presentations ‘go beyond themselves’ (A197/B242). At the very least, this claim strikes me as sufficiently interesting and plausible to pursue its implications, a topic which the remainder of this thesis is primarily orientated towards (though some objections are considered in Section 5.3).

Before turning to this however, there is one final observation regarding the rule-subjectingness of experience which warrants attention. To this point, I have limited myself to describing rules in terms of *commitments* holding of the subjects and certain beliefs. Nevertheless, we could speak just as easily of requirements, prohibitions, permissions (Landy, 2009: 18). Importantly, these normative relations are all alike in being intrinsically universalisable and inter-subjective. To grasp the idea of being committed to something, it is necessary to grasp the universalisability of the commitment: I can recognise that I am committed to something only if I recognise that any rational being in relevantly equivalent circumstances would also be so committed. The same holds, *mutatis mutandis*, for prohibitions, requirements etc. Consequently, insofar as I take other rational creatures to be in relevantly equivalent circumstances, I am obliged to treat a commitment of mine to be a commitment of theirs also. We insist that if they are valid for us, they are valid “as well for everyone else” (Prol: §18, 45). As such, these normative relations are inherently inter-subjective. Insofar as objective experience is essentially rule-subjecting, it follows that it implies inter-subjectivity and universalisability:

“Objective validity and necessary universal validity (for everyone) are exchangeable notions.” (Prol: §19, 45)

The intersubjective nature of rules is significant because it serves as a very useful *test* for whether something is rule-subjecting. That is, it constitutes a criterion by which we can determine whether something is rule-subjecting or not. Roughly, we can ask, is it such that it commits the subject to requiring certain things of *others*? Does it, in this sense, exhibit universal validity? For example, it is possible to determine that the objective state ‘the boat is floating downstream’ is rule-subjecting because the subject is obliged to demand that others must agree. In other words, the intersubjective nature of rules is a criterion by which we can

establish whether something is rule-subjecting or not. This will be important in the next section.

## 5.2. *The priority principle*

Any adequate discussion of Kantian concepts must make reference to judgement, for one of Kant's most well known theses is that judgement takes priority over concepts:

“...the only use that the understanding can make of these concepts is to judge by means of them” (A68/B93)

“[Concepts are] predicates of possible judgements” (A69/B94)

However, while commentators agree that judgement is prior to concepts for Kant, there is no agreement on how this ‘priority’ ought to be understood. Heis (forthcoming) lists eight distinct interpretations of the principle which have been suggested in the literature. In my view, all of these interpretations either fail on exegetical grounds, or are inadequate to the task at hand: that of substantiating the relevant notion of *priority*. However, rather than engage in the lengthy undertaking of refuting each of these in turn, I shall instead simply strike out with my own account which is, to my knowledge, original and so stands or falls on its own merit.

In order to understand the priority principle, it is first necessary to understand what Kant means by ‘judgement’. Unfortunately, this presents a challenge in its own right for Kant offers numerous definitions of judgement. Although Longuenesse (1998: 81) urges that these definitions can be shown to be equivalent, the proofs she offers of this are not at all trivial. Rather than engaging in a comprehensive review of all of Kant's definitions of judgement then, I want instead to underline what is at least a central element of his conception of judgement: their conceptual composition.

“In every judgement the given concepts can be called logical matter (for the judgement).” (A266/B322)

“...thus in the judgement ‘all bodies are divisible’, *the concept* of the divisible... is here applied in particular to *the concept* of body” (A68-69/B93-94) [my emphasis]

As Longuenesse (1998) emphasises, for Kant all judgements involve the combination or relation of concepts (see also, Hanna 2001: 60-63). This point is also underlined by an inspection of Kant's table of the forms of judgement. Although I do not discuss this table fully until Section 5.4, suffice to say that according to Kant, all judgements must be composed of at least two concepts since all judgements are either categorical, disjunctive or hypothetical in form (A70/B95).

The mere fact that judgements are conceptually constituted does not have any direct implications for the priority principle. In particular, the mere fact that judgements are composed of concepts does not entail that they are prior, in any substantive sense, to concepts. A pad of paper is composed of individual sheets, but that does not render the pad logically prior to the sheets. Nevertheless, I do think that the conceptual constitution of judgements suggests a conjecture which, if shown to be correct, would substantiate the priority principle. The conjecture is as follows: what if concepts only exist as the constituents or components of a composite whole. That is, what if concepts are only concepts insofar as they have been combined or composed with other concepts. Call this the composition conjecture. Recall that, for Kant concepts are rules, so the conjecture here is that rules only exist insofar as they have been composed with other rules. Were this conjecture correct it is fairly clear how this might substantiate the priority principle. Basically, since judgements just are compositions of concepts, if concepts only exist in such compositions, this would raise the possibility of concepts only existing as the constituents of judgements. This certainly would amount to the priority of judgement over concepts. This is clear enough to warrant an investigation of the composition conjecture.

In my view, there is good reason to think that the composition conjecture is correct. Consider the following judgements:

- Birds fly
- All bodies are divisible

As Macfarlane (2002: 36) says, despite these judgements not being grammatically prescriptive, i.e. they are not imperatival, they nevertheless “imply prescriptions”. To judge that birds fly is to be committed to certain beliefs, and as such, such compositions of concepts exist in “a network of various *normative* relations: commitments, prohibitions, permissions, etc” (Landy, 2009: 18). For example, ‘all bodies are divisible’ commits one to the belief that ‘there is nothing which is a body and is not divisible’. Again, this is not a prediction or



hypothesis about the flow of one's psychological states. Rather, it establishes a web of obligations. It establishes what one is correct in doing and incorrect in doing, and so renders the subject responsible or assessable in certain ways. As such, it is perfectly natural to speak of the constituents of these judgements as *rules*. For example, substituting 'lions' for 'bodies' in the composition 'All bodies are divisible', commits the subject very differently. It follows that in each case, whoever so judges is subject to distinct rules.

Furthermore, these judgements are valid universally and necessarily. If I judge that 'all bodies are divisible', I require that other peoples' judgement be consistent with my own (Heis, forthcoming: 20). Crucially, this does not entail that other people must assert 'all bodies are divisible', but simply that they must not contradict it. I am committed to other rational creatures finding things as I find them, and where this is not the case, I am committed to seeking reasons for this divergence. In this sense, judgements are inter-subjective and universalisable. By contrast, consider the constituents of the above judgements in isolation:

- 'Bird'
- 'Body'

, or a list of these constituents:

- {Body, divisible, bird, fly}

In neither case of thinking these constituents is the subject committed to anything. For example, just because I think at one moment {body, divisible, bird, fly}, this does not render the subsequent thought {mind, simple, mammal, swim} incorrect or a mistake. The mere fact of thinking such representations does not demand anything of the subject (Corr: Letter to Beck, July 1792, 420). Nor is there any sense to be made of the idea that these representations prescribe anything which is valid universally and necessarily. There is nothing inter-subjective about thinking 'bird': such thinking does not compel the subject to, for example, demand that others think alike. Indeed, the idea is faintly absurd. What this suggests is that rules only exist in the context of, or as constituents of, judgements: it suggests that the composition conjecture is correct.

Importantly, it is not the case that all compositions of concepts are such that the constituent concepts are rules. For example:

- Red bus

- Bus red

In these examples, the concepts ‘bus’ and ‘red’ are hierarchically composed, as indicated by the asymmetries in their meanings: a red bus is a bus that is red whereas bus red is a red characteristic of buses. Nevertheless, despite being compositions in this sense, these representations do not commit the subject to anything anymore than the uncomposed examples above. As such, there are no rules here. A consequence of this qualification is that judgement needs to be understood more precisely as *those combinations of concepts in which the concepts are rules*. I think this is what Kant has in mind when he says:

“I have never been able to accept the interpretation which logicians give of judgement in general. It is, they declare, the representation of a relation between two concepts. I do not here dispute with them as to what is defective in this interpretation-that in any case it applies only to categorical, not to hypothetical and disjunctive judgements... I need only to point out that the definition does not determine in what the asserted relations consists.” (B141)

Kant does not deny that judgements are compositions of concepts here, but insists that this is not specific enough. A more precise definition would be that judgements are those compositions of concepts in which the concepts are *rules*. Judgements are the minimal units of responsibility. Brandom (2009: 32-33) sums the picture up well:

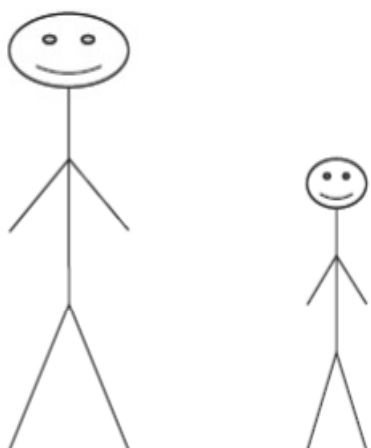
“What distinguishes judging and intentional doing from the activities of non-sapient creatures is not that they involve some special sort of mental processes, but that they are things knowers and agents are in a distinctive way *responsible* for. Judging and acting involve *commitments*. They are *endorsements*, exercises of *authority*.

Responsibility, commitment, endorsement, authority-these are all *normative* notions.”

Insofar as the composition conjecture has been established, a strong version of the priority principle follows. Quite simply, judgements are prior to concepts insofar as concepts only exist in the context of a judgement.

This discussion implies that when I spoke in the previous section of concepts as rules, I was oversimplifying the matter in a crucial respect. Namely, I was eliding the fact that rules only arise in the context of a *judgement*. All concepts, insofar as they are rules, are constituents of judgements. For example, consider the sensory array:

Figure 7.



As has been discussed, in order to amount to perceptual experience, such a deliverance of the sensibility must be subsumed under rule(s). However, since rules only exist in the context of a judgement, this subsumption must necessarily involve a judgement. For example, judging that the tall man is next to the small man would amount to bringing this manifold under rules. Yet no fragment of this judgement would suffice in this respect. This is why Kant calls the power of judgement “the faculty of *subsuming* under rules” (A132/B171). Sensory representations are brought under rules by means of judgement (Allison, 2008: 8; Davidson, 2004: 138).

A terminological complication follows from this priority principle. Notice that in the above cases where the constituents of judgements are isolated or listed in abstraction from any judgemental context, it is natural to say that what are being so isolated are *concepts*. Indeed, I have spoken of ‘uncomposed concepts’ and compositions of concepts which fall short of a judgement e.g. ‘bus red’. Nowadays, concepts are generally supposed to have their identity conditions irrespective of their composing any judgemental wholes. Indeed, the idea of composition implies as much. For example, a chair may be composed of various parts; yet, the existence of these parts does not depend upon their entering into the mereological relations constitutive of a chair. Despite this however, it should be clear that if the priority principle, that rules only exist in the context of a judgement, is correct, talking of concepts in abstraction from judgemental wholes is, strictly speaking, precluded on Kant’s account. This is because, for Kant, concepts just are rules. If a concept does not compose a judgement then it is not a concept.

This is not inconsistent with viewing judgements as composed of concepts, but it does necessitate an alternative conception of composition. Kant distinguishes between two kinds of mereological involvement: constituents of a whole are either juxtaposed (*Zusammensetzung*) or unified (*Verknüpfung*) (B201). As Bell (2001) underlines, in the case of juxtaposition, the relations constituents stand in are external or contingent in the sense that neither the identity nor the existence of the constituents depends in any way upon the relations. By contrast, in the case of unity, the relations constituents stand in to each other are essential to their identity and/or existence. In the former case, the whole is merely an aggregation of the parts whereas in the latter, the parts are abstractions from the whole. The mereological relations constitutive of judgement are those of unity, not juxtaposition; in a judgement, concepts compose a unity (Bell, 2001: 11). Consequently, concepts are abstractions from judgements.

“There is no description of what it is for a concept, as an entity, to have the capability of representing things mediately by itself. For Kant these are the capabilities not of concepts but of understanding, capabilities exercised in acts of judgement. Kantian concepts are best viewed not as entities which represent on their own, but as abstractions from these cognitive capabilities exercised in judgement: they are rule-like modes by which acts of judgement can be made.” (Pereboom, 1988: 328)

There is an extremely significant consequence of Kant’s priority principle: it entails that the capacity for judgement is central to the possibility of experience. In the previous section I outlined how, according to Kant, all experience is essentially rule-subjecting. In this section, I have argued that the priority principle should be understood as the claim that all rules are possible only as constituents of a judgement. It follows from this that experience is essentially judgemental. In other words, in the context of Kant’s transcendental investigation, a consequence of the priority principle is that it moves us from the claim that experience is essentially normative to the claim that experience is essentially judgemental.

1. Objective experience requires concepts.
2. Concepts only exist as the constituents of judgements.
3. Therefore, all objective experience is judgemental.

This argument chimes with some of Kant’s most famous pronouncements regarding judgement. Recall that I earlier noted that Kant offers numerous definitions of judgement.

Some of these refer to the conceptual composition of judgements which I have already discussed. Others however, define judgement in *functional* terms.

“A judgement... [is] a relation that is objectively valid, and so can be adequately distinguished from a relation of these same representations in which there would be only subjective validity, e.g. in accordance with laws of association.” (B142)

“[Judgements are] act[s] by which given representations first become cognitions of an object” (Meta: 145-6, fn. 3)

Anderson (2008: 102) sums up this function well when he says:

“Stripped of Kantian jargon, the thought seems to be that a judgement, by contrast to a mere psychological association of ideas, relates its constituent representations in such a way that they represent things as being a certain way.”

Taken in isolation, the ascription of this function to judgements appears puzzling. After all, this ascription is clearly non-trivial and *prima facie*, it might be wondered what justifies the claim that judgement is so integral to relation to an object?<sup>101</sup> Given the foregoing account of the priority principle however, a justification is readily available: concepts are integral to relation to an object; judgement takes priority over concepts; therefore, judgement is integral to relation to an object. Moreover, that the priority principle is capable of rationalising what appears to be a distinct Kantian thesis – the connection between judgement and relation an object – in turn lends credibility to the interpretation of the priority principle I have defended.

In light of this, it is possible to extend the earlier noted identification of the faculties of the understanding, of concepts and of rules:

“We can trace all acts of the understanding back to judgements, so that the understanding in general can be represented as the capacity to judge.” (A69/B94)

So:

Faculty of the understanding = Faculty of concepts = Faculty of rules = Faculty of judgement

---

<sup>101</sup> Indeed, although this connection between objectivity and judgement is accorded a position of considerable prominence in much of the secondary literature, I know of no *explanation* of it.

### 5.3. *Objecting to Conceptualism: concepts and schemata*

The foregoing account of experience is a version of Conceptualism: the view that all intentional content requires the application of concepts.<sup>102</sup> However, Conceptualism is not without its critics, and the position has been the topic of considerable debate in recent years (Ginsborg, 2008; de Sa Pereira, 2013). A major problem for Conceptualism is that it seems to over-intellectualise experience. Is it really plausible that all experience – any mental state bearing intentional content – requires an intellectual apparatus as sophisticated as that of concepts? This problem is surely most troubling with respect to perceptual experience, and many commentators have argued it is possible to explain the nature of perception without invoking concepts (e.g. Dreyfus, 2007; Hinzen and Sheehan, 2013). While somewhat compelling, this rather sweeping problem is closely related to a more specific objection which concerns both animal and infant cognition. Both animals and infants, so goes the objection, enjoy perceptual experience: they encounter the world perceptually. Given this, the conceptualist is compelled to ascribe concepts to them. However, especially when the normative and holistic character of concepts is underlined, along with their essential dependence on the capacity for judgement, this ascription does not appear viable (see, Hanna, 2011: 330-1). We want to be able to grant animals and infants perceptual experience without burdening them with such a rich intellectual capacity, but this entails the renunciation of Conceptualism.

In this section, I address these worries, even if not dispelling them altogether, by drawing upon the first chapter of the *Analytic of Principles*, the *Schematism*. I argue that although Kant certainly is a conceptualist, he has the philosophical resources to both do justice to the broad ‘anti-conceptualist intuition’ with respect to perception – that concepts are not needed to explain perception - and more importantly, meet the specific challenge from infant and animal cognition.

In the *Schematism*, Kant is occupied with the issue of concept application, or equivalently, the subsumption of intuitions under concepts. Specifically, Kant thinks that there is a problem regarding how it is that sensory states are *recognised as instances of*

---

<sup>102</sup> That Kant is a conceptualist is the dominant view in the literature. Indeed, expositions of Conceptualism commonly take Kant’s observations in CPR and elsewhere as a point of departure, treating him as perhaps the first major advocate of the thesis (see, e.g. McDowell, 1994). Admittedly, this exegetical view is not universally accepted. Hanna (2005) argues that Kant can be understood as the father of Non-conceptualism also! My own view is clearly that Kant is a conceptualist – this is entailed by the account I have attributed to him. My concern in this section then is not with the exegetical question of Kant’s Conceptualism but with its philosophical credibility.

concepts: how the former are subsumed under the latter or how the latter are applied to the former. One way of understanding Kant's worry here is to recall his discursivity thesis: experience involves two irreducible and qualitatively distinct stems, the understanding and the sensibility. While emphasising the distinctness of these two faculties has its advantages, a puzzle inevitably arises with respect to their interaction. In particular, how is it that concepts, as the denizens of the understanding, can be known to apply to intuitions which, as products of sensibility, play a wholly distinct cognitive function? How could concepts, which are qualitatively non-sensible, specify the sensible conditions of their realisation? Quite generally, concepts, as general functions, seem far too abstract to determine whether particular sensory states count as instances. This is a worry Kant raises as early as the *A Deduction*, hinting there also at a solution:

“The two extremes, namely sensibility and understanding, must stand in necessary connection with each other through the mediation of this transcendental function of the imagination.” (A124)

This transcendental function of the imagination Kant calls a *figurative synthesis*, or *synthesis speciosa*. Although at this point in the *A Deduction* it is not yet clear in exactly what sense the understanding and sensibility require ‘mediation’, the suggestion is that there must be a mediation and that this is supplied by an act of imaginative synthesis. The issue of mediation receives its full hearing in the *Schematism*, where this is explicated as the challenge of how concepts are applied to intuitions.<sup>103</sup> Here, the figurative synthesis of the imagination is called a schema. It is by being connected to a specific schema, that concepts are known to apply to intuitions: schemata are the basis of all perceptual use of concepts, whether they be pure sensible (i.e. geometrical) concepts, pure intellectual concepts or empirical concepts (A140/B180).<sup>104</sup>

“Even less is an object of experience... ever adequate to the empirical concept; rather, that concept always refers directly to the schema of the imagination, this schema

---

<sup>103</sup> For the connection between the *A Deduction* and the *Schematism*, see Hanna (2001: 53) and Longuenesse (1998: 116).

<sup>104</sup> It has been contended that Kant's problem in the *Schematism* involves only the application of pure concepts, i.e. categories, and excludes the application of empirical concepts since these are sufficiently ‘homogenous’ with sensible intuition so as to require no intervening schema (Guyer, 1987). In places Kant certainly suggests as much. Nevertheless, I agree with Pendelbury (1995) that the balance of assertions in the *Schematism* makes it clear that the problem extends also to pure sensible and empirical concepts. The application of the categories is simply a special or particularly challenging instance, of a more general problem regarding concept application *tout court*.

being a rule determining our intuition in accordance with such and such a general concept. The concept *dog* signifies a rule according to which my imagination can delineate universally the figure of a four footed animal without being limited to any particular figure, such as experience or any possible image which I can represent *in concreto*, actually presents.” (A141/B180)<sup>105</sup>

Basically, the thought is that we know how to apply concepts to intuitions, only if the latter have already been configured by an act of synthesis performed by the imagination: a schema. Now, there is perhaps no idea which Kant utilises more persistently in CPR than that of synthesis. Indeed, given the profusion of syntheses in CPR, coupled with Kant’s assertive prose and eschewal of examples, explicating a schema as a kind of synthesis is liable to obscure rather than clarify matters. At the limit however, a synthesis, for Kant, denotes simply the combination of a manifold in some determinate way.

“Now since every appearance contains a manifold, and since different perceptions therefore occur in the mind separately and singly, a combination of them, such as they cannot have in itself, is demanded. There must exist in us an active faculty for the synthesis of this manifold. To this faculty, I give the title, imagination.” (A120)

Schemata combine manifolds in some determinate fashion: group stimuli according to certain templates. One never encounters a pure un-discriminated perceptual manifold. Rather, a manifold is always organised or differentiated into distinct qualities of varying salience and significance (Hanna, 2001: 52). In this sense, every manifold is riddled with internal contours which cut either one way or another; even our most primitive confrontations of the world exhibit a structure of this kind. According to Kant, this is the result of a mere sensory manifold being taken up and synthesised according to schemata. Most fundamentally then, a schema is simply a capacity for sensory discrimination.

It follows from this that a schema is also a capacity for recognising generic identities and differences across *distinct* manifolds. If a schema combines a manifold,  $m_1$ , as *so* and *so*, and combines at a later time, a distinct manifold,  $m_2$ , as *so* and *so*, this amounts to the (partial)

---

<sup>105</sup> There is a puzzle regarding this quote. Although from the context it is clear that Kant means to assert that empirical concepts, such as ‘dog’, require schemata for their application, it is odd that he says that the concept of dog “signifies” an imaginative synthesis, for this synthesis is a schema, not a concept. Allison (2004: 208) contends that Kant here simply “misstates his own position, referring to the concept of dog, when he clearly means the *schema*”. Alternatively, Kant statement here might be interpreted as saying that the concept ‘dog’ signifies, in the sense of being linked or connected to, a schema of the imagination (this seems to be Pendelbury’s view (1995: 789)).



identification of the respective manifolds. For example, if  $m_1$  is synthesised according to the schema, ‘dog’, and so is  $m_2$ , the manifolds share a certain affinity: namely they have both been organised according to the schema of ‘dog’. In other words, in synthesising *particular* manifolds a schema amounts to a capacity for recognising generic identities and differences across *distinct* manifolds,<sup>106</sup> and so serves to establish similarity classes of manifolds. As Longuenesse (1998: 62) puts it, schemata yield ‘series’ of intuitions.

Having ascribed these functions to schemata, it is, I think, relatively easy to state how being connected to a particular schema explains how concepts can be known to apply to intuitions. Essentially, a concept applies to an intuition insofar as that intuition has already been synthesised according to the schema with which that concept is connected. The concept, ‘dog’, applies to an intuition if that intuition has been synthesised according to the schema, ‘dog’. This explains how intuitions are subsumed under concepts. In this sense, from the perspective of the problem of concept application, schemata do all of the ‘heavy lifting’: concepts just need to be in a one-to-one relation with them and the conditions of their application are secure (Longuenesse, 1998).

Although this illuminates the function of schemata, it does little to specify their categorical nature. Kant emphasises that the schema of a concept is not simply an image, though schemata may yield images (A140-1/B179-180). The reason for this seems to be that images are too specific to specify the sensible conditions of concept application. For example, any image of a dog will necessarily be specified with respect to certain qualities, e.g. colour, size, which will inevitably exclude certain dogs from its range.<sup>107</sup> Instead of images, Hanna (2001: 52) argues that schemata correspond to what are known as prototypes or exemplars in contemporary literature. However, while Kant does describe examples as “leading strings” (A134-5/B174), he does not equate schemata with prototypical exemplars of the sort Hanna suggest (cf. Allison, 2004: 205-208). Sellars (1978) suggests that to synthesise a sensory array according to a schema is to token a complex demonstrative phrase e.g. ‘this brick with a red and rectangular facing surface’. There is no predication and no judgement here. Instead, schemata yield subject terms for these judgements. This view is also problematic however, insofar as it portrays schemata as ‘quasi-conceptual’ (Allison, 2004: 188). The notion of a subject is a conceptual matter if anything is, so to say that schemata yield subjects for

<sup>106</sup> The notion of ‘recognition’ relevant here is evidently not, or at least need not be, a conscious one.

<sup>107</sup> This echoes Hume’s own observations regarding the impossibility of an image which is not specified with respect to certain qualities. The difference however, is that since Hume saw no alternative to the view that ideas are images, he was forced to adopt a highly convoluted account of abstract ideas (see fn. 24).

judgements is to say that they are, in and of themselves, a kind of conceptual operation. It is difficult to reconcile this with the stated function of schemata as enabling concept application: insofar as concept application is a problem, it cannot be solved by postulating ‘another’ class of conceptual operations. I agree with Pendelbury (1995: 787) that it is best to interpret schemata as dispositional properties. Insofar as schemata serve to classify or discriminate our intuitions according to certain similarity classes, this can be understood dispositionally: dispositions to *respond* to stimuli in similar ways. On this view, to recognise the general in the particular then, is to be disposed to react to the particular in a manner similarly to how one reacts to other particulars of the same kind.

Ultimately perhaps, the precise categorical character of schemata is not too important. After all, as Pendelbury (1995: 786, fn. 18) points out, within a few pages (A140-7/B179-186), Kant describes schemata variously as *representations, rules, products of the imagination, syntheses, phenomena of agreement* and *sensible concepts*. So it is unlikely that, from an exegetical perspective at least, a conclusive view of what schemata are is possible. Instead, what really matters is the function or role schemata play. As I have described it, schemata actually perform a three-fold function:

- (a) Synthesise manifolds in determinate ways.
- (b) Establish similarity (and non-similarity) classes of manifolds.
- (c) Specify the conditions of concept application.

Notice that while (a) and (b) are inextricable from each other, they could, in principle at least, be detached from (c). This is important for what follows.

Even as compared with other equally difficult and relatively obscure sections of CPR, the status and reputation of the *Schematism* amongst commentators is especially contentious. As Allison (2004: 202) notes, the chapter is unusual in being sometimes viewed as completely superfluous. For example, Warnock (1948) dismisses the *Schematism* as altogether misconceived, arguing that Kant’s “peculiar language” lures him into pursuing a nonsensical problem. The charge against the *Schematism* is that the central issue upon which the chapter turns, how concepts are applied, is simply confused. In asking how it is that concepts are applied or how intuitions are subsumed under concepts, the *Schematism* is premised upon the separation of concept possession from concept application: there can be no ‘problem’ of concept application if possession of a concept is, at least in part, constituted by knowing how to apply it. The charge against the *Schematism* is that this premise is false: no

such separation of concept possession and application is possible. To have or possess a concept *is* to know how to apply it, or in Kant's terms, concepts are their own schemata. This is why, according to Warnock, Kant is pursuing a nonsensical problem (for a related view, see Bennett, 1966). After all, it is only on the basis of concept use that we ascribe concept possession to others. If someone was unable to correctly sort a collection of dogs and sweet wrappers into dogs and non-dogs, we would certainly not attribute to her the concept of a dog. Indeed, it is far from clear that we can even conceive of what it would be like to possess the concept of a dog, without being able to recognise instances of dogs (and therefore, also, non-dogs).

This Warnockian critique is far less persuasive however, if we turn from commonplace concepts such as 'dog', to less familiar ones. For example, consider the concepts of 'tadpole' and 'bone marrow'. I certainly know some things about tadpoles and bone marrow. Tadpoles are the larval form of frogs, and bone marrow is the kind of thing which can be transplanted in an operation and produces red blood cells. In this sense, I possess these concepts, even if they are relatively impoverished. However, I would be completely hopeless at recognising either tadpoles or bone marrow. While I would have little trouble classifying some things as not-tadpoles, e.g. book cases and elephants, there would be a good number of non-tadpoles I would fail to distinguish from tadpoles proper, similarly with bone marrow. For less familiar concepts, it seems that we can separate concept possession from knowing how to apply it. Ironically, the Warnockian critique of the *Schematism* stems largely from the inadequacy of Kant's own 'dog' example. By focusing instead upon less familiar concepts, the distinction between concept possession and application is apparent.

Still though, it might be questioned just what this establishes. It may be admitted that it is possible to conceive of possessing a concept without being able to apply it. Nevertheless, why does this necessitate the postulation of a 'third thing' mediating between concepts and intuitions? That is, even if we admit the coherence of the problem of the *Schematism*, in what sense is the postulation of schemata a solution to it? True, as syntheses of the imagination, schemata are supposed to be 'both' intellectual and sensible (A138/B177), but it unclear how much philosophical weight this claim can withstand. More importantly, rather than addressing the problem of the *Schematism*, the postulation of schemata, at best, seems merely to re-locate it. The problem of concept application was how to determine whether a concept applies to a particular sensory state. The proposed solution is to say that every concept is

connected to a schema which determines whether it [the relevant concept] applies to a particular manifold or not depending upon whether the manifold has been synthesised according to that schema. Yet, far from being a true solution to the problem, it seems that this simply shifts the problem from concepts onto schemata. That is, the problem of concept application can be re-formulated as the problem of schemata application. How do we determine that a particular manifold falls under one schema rather than another? Insofar as this is a problem for concepts, why shouldn't it be a problem for schemata also? Perhaps the difference has to do with the different categorical nature of concepts and schemata or perhaps schemata are devoid of the kind of answerability to particulars which concepts exhibit. However, such a suggestion is difficult to substantiate given Kant's vagueness on the types of things schemata are. It does seem that once we admit that there is a problem of concept application, it is hard to deny that the postulation of schemata is not afflicted by the same problem. As such, even granting that the problem of the *Schematism* is coherent, it is far from clear that the postulation of schemata can solve it (Bennett 1966: 150).

I think that this objection to the *Schematism* is plausible. Consequently, were the distinction between concepts and schemata motivated solely on the basis of the problem of the *Schematism*, it would, at least arguably, not be justified. Equivalently, were the function of schemata exhausted by or limited to that of determining concept application, their postulation would stand on thin epistemic ground. Crucially however, - and it is at this point that we need to move away from the *Schematism* - this is not the case. It is possible and necessary to speak of schemata in abstraction from the specific problem of concept application; they play a much broader role in Kant's philosophy. Firstly, notice that there is no incoherence in possessing schemata in the absence of any corresponding concept. For example, Travis (Ms.: 17) notes that of the two following questions it is possible to be able to answer (1) without being able to answer (2):

1. 'Which things (hereabouts) are pigs?'
2. 'What cases of something's being as it is would be (count as) a case of something being a pig?'

The generalisation, 'listen for the grunters' is an answer to (1); it serves to discriminate pigs from non-pigs.<sup>108</sup> Yet, imparting such a schema to someone unfamiliar with pigs would not

---

<sup>108</sup> That such a schema would occasionally lead one to inaccurately discriminate pigs, i.e. there clearly could be pigs that do not grunt, is irrelevant since the notion of accuracy is itself conceptual.

provide them with the means to answer question (2): one could have the means for recognising certain generic identities or similarities – namely, the similarity of everything which grunts - without having the concept of ‘pig’, or knowing anything about pigs. Put this point another way. When Pia recognises a pig by sight, she is presumably sensitive to the presence of certain visual features. This sensitivity amounts to a perceptual schema of pigs insofar as it provides Pia with the ability to recognise that the distinct perceptual episodes, X and Y, resemble each other and so can be grouped together. However, such a sensitivity does not amount to the concept of a pig. Pia may be able to discriminate X and Y, on the one hand, from Z, on the other, without knowing what X and Y are: without possessing any concept which X and Y fall under. This is what Pendelbury (1995: 787) driving at when he says:

“[Schemata] yield discriminating responses rather than articulate, classificatory judgements; and significantly, they do not involve the capacity to make analytic judgements, like *Every triangle is a closed figure* and *No completely green square is red*, on the basis of reason, or the capacity to make synthetic judgements, like *Some red things are not squares* and *No horse is a cow*, on the basis of experience. There is therefore good reason to regard schemata as more primitive than concepts.”

In other words, concepts and schemata are doubly dissociable. Not only is it possible to possess concepts without possessing corresponding schemata, it is also possible to possess schemata without possessing any corresponding concepts (cf. Davidson, 2004). This is very significant. If it is possible to possess schemata in the absence of any corresponding concept, then the postulation of schemata can be detached from the issue of concept application: their postulation need not be justified by recourse to this problem. Rather, it is possible for their postulation to be justified on the basis of non-conceptual work they perform. In other words, this confirms what I hinted at earlier, that of the three-fold function schemata perform for Kant: (a) synthesising manifolds in determinate ways, (b) establishing similarity (and non-similarity) classes of manifolds, and (c) specifying the sensible conditions of concept application, it is possible to isolate the first two functions, which make no essential reference to concept application, from the third, which does.<sup>109</sup>

Indeed, we find Kant employing schemata to perform non-conceptual, philosophical work. This shows up in the three-fold synthesis of the *A Deduction* where Kant is undertaking

---

<sup>109</sup> As I outline, I think that there is good evidence that this was Kant’s view. At the limit however, my claim is that the distinction between concepts and schemata provides Kant with the resources to adopt this view and that he *should do so*.

his ‘encirclement’ of the empiricist position. Kant begins by noting that it is an “empirical law” that representations which are constantly conjoined become associated (A100), a fact well familiar to readers of Hume. He then suggests however, that this association presupposes certain conditions.

“Suppose that cinnabar were now red, then black, now light, then heavy; or that a human being were changed now into this and then that animal shape; or that on the longest day of the year the land were covered now with fruit, then with snow and ice. In that case my empirical imagination could not even get the opportunity, when presenting red colour, to come to think of heavy cinnabar.” (A100-1)

This seems straightforward enough: were properties never constantly conjoined, but instead conjoined only randomly, our powers of association would never be exercised, remaining concealed as a “dead and unknown faculty” (A100). However, this constant conjunction itself presupposes certain conditions. That is, association, in presupposing constant conjunction, presupposes also the power of recognising generic identities and differences across distinct manifolds. Roughly, if A and B are to be constantly conjoined, it is necessary to be capable of discriminating across distinct manifolds As and Bs. There must, in some sense, be recognition that any particular A is the same as As which have gone before and *mutatis mutandis* for any particular B. Were this not the case, every conjunction of A and B would be a ‘new’ conjunction, and there would be no *constancy* to speak of.<sup>110</sup> This is clear enough and, for Kant, this power is, of course, a type of synthesis, a schema:

“The generic identity of the empirical intuitions reproduced by associative imagination is shown to be the generic identity of the *acts of successive synthesis* which in turn generates each of these particular empirical intuitions “as” specific unities of a manifold.” (Longuenesse, 1998: 47)

It should be clear that schemata are doing ‘non-conceptual’ work here. After all, association is an empirical law which holds also of animal psychology. However, animals lack concepts and so while association presupposes schemata, it cannot presuppose concepts.

Separating schemata and concepts in this way seems to be plausible. Indeed, interestingly, it has a contemporary analogue, if not an outright correspondence, to the one

---

<sup>110</sup> This is far from being a complete account of the notoriously nebulous three-fold synthesis of the *A Deduction*. Nevertheless, it serves to show that schemata lie at the heart of it (see, for discussion, Longuenesse, 1998: 116, fn. 29).

drawn by Hinzen and Sheehan (2013: Chpt. 2) between percepts and concepts. A percept is understood to play the same role as a Kantian schema: it serves to unify “a mass of environmental stimuli” (2013: 78) according to certain classifications: a dog’s skin, a piano playing, etc. It groups stimuli according to such abstract classes.<sup>111</sup> However, percepts do not yield experience; they are non-intentional; they do not move in the space of reasons, and the question of truth and falsity (and error) does not arise for them.<sup>112</sup> Instead, all of this presupposes the possession of concepts which are understood as distinct, though related, intellectual capacities: intentionality, intensionality and concepts “arise together” (2013: 82).<sup>113</sup>

The reason why this distinction of concepts and schemata is important, indeed, the reason why I have gone such lengths to develop it is that it serves to defuse the objections to Conceptualism outlined earlier. Consider the objection from animal and infant cognition; the distinction of schemata and concepts affords a response here. Specifically, although animals and infants lack concepts, they can still possess schemata. Indeed, this is exactly how Kant characterises the matter:

“The door is something that does, it is true, belong to the stall and can serve as a characteristic of it. But only the being who forms the judgement: *this door belongs to this stable* has a distinct concept of the building, and that is certainly beyond the power of animals.” (FS: §6, 103)

Kant grants animals the capacity to distinguish amongst their representations and recognise generic identities and differences. Although this statement is found in a pre-critical (1762) text, Kant asserts the same point in *Logic* (Intro., VIII, 71) where he distinguishes being:

- (1) “...acquainted with something, or to represent something in comparison to other things, both as to sameness and to difference”

---

<sup>111</sup> Hinzen and Sheehan claim that pre-conceptual perceptual categories have a *kind of* semantics. I would prefer to say that there is no semantics here since there is no intentionality (indeed, Hinzen and Sheehan themselves want to deny percepts intentionality). Instead of ascribing percepts a semantics, it is better to understand perceptual categories dispositionally, in the same way which I have suggested schemata ought to be understood.

<sup>112</sup> Hinzen and Sheehan do not postulate, at least not explicitly, any relation of dependence between truth-evaluability and moving in the space of reasons. Nevertheless, as should be clear from the earlier sections of this chapter, for Kant these are intimately related.

<sup>113</sup> This is a little misleading insofar as I have left out what, for Hinzen and Sheehan, is the key component: grammar. They argue that intentionality etc. arises only with *grammar*. Although I agree, I have left this element out for now since it evidently goes beyond Kant’s own position. Nevertheless, Chapter seven of this thesis is orientated towards showing how Kant’s position can be updated or augmented along these lines.

and,

- (2) "...understand[ing] something i.e. to cognize something through the understanding by means of concepts"

Kant asserts that (1) applies to animals, but (2) does not: they make sensory discriminations but do not cognize objects. In other words, they possess schemata, but not concepts.

Interestingly, Davidson adopts precisely this view of animals:

"A creature does not have a concept of a cat merely because it can distinguish cats from other things in its environment. For all I know, mice are very good at telling cats apart from trees, lions and snakes. But being able to discriminate cats is not the same as having the concept of a cat" (Davidson, 1999: 8; cf. Davidson, 2004)

This distinction between discriminatory powers and conceptual powers vitiates the objection from animal cognition. Just because animals lack concepts, it is not the case that Kant is obliged to regard their mental lives as a 'blooming, buzzing confusion'. On the contrary, animals can be seen to possess sophisticated mental powers, schemata, enabling them to categorise their environment and navigate it ably. In other words, it is possible to account for the rich encounters animals appear to have with the world without ascribing them concepts. Of course, this thesis does still assert a radical discontinuity between the (non-conceptual) mental powers of animals and those of humans. Despite ascribing animals discriminatory capacities, Kant must deny that they enjoy experience, for this presupposes concepts. Perhaps the non-conceptualist may insist that this is unacceptable: animals possess not only discriminatory capacities but have fully intentional experience. But now the objection looks far less forceful and it is unclear why the conceptualist needs to accept this. On the contrary, although I certainly do not want to undertake a review of the comparative cognition literature here, it is worth noting that the cognitive discontinuity of humans and animals has been prominently defended in recent years on the basis of empirical considerations (Penn et al, 2008; see Section 7.3 for discussion).

This response can be extended without complication to infants also. That is, while it is certainly true that infants enjoy sensory encounters with the world, plausibly this can be construed in wholly schematic terms. For example, it is well established that pre-linguistic infants discriminate amongst objects according to abstract kinds such as 'box', 'car', 'toy', 'cat' (Xu, 2005). Yet, insofar as ascribing schemata to infants is potentially sufficient to



account for the ways in which they interact with their environment, Kant can deny that they have experience and therefore need not ascribe them concepts.

As well as the animal and infant objection, at the outset of this section I mentioned also an anti-conceptualist ‘intuition’ regarding perception: namely that it can be accounted for without recourse to concepts. I call this an ‘intuition’ rather than an objection proper principally because it is not, at least as stated, an argument at all. Rather it is just the assertion of a non-conceptualist intuition and so is unlikely to impress the conceptualist. Nevertheless, even here I think that Kant’s distinction of concepts and schemata can go some way towards addressing, and even doing justice to this intuition. According to Kant, concepts are known to apply to particular intuitions via the mediation of schemata. However, as I have shown, this pre-conceptual figurative synthesis of a manifold can be isolated and separated from its role in concept application in both developmental terms and comparative terms. Developmentally and comparatively, it is possible to distinguish the sensory discrimination and apprehension which is secured by schemata, from experience, which presupposes also concepts. Given this however, - and this is the crucial point - why think that in the case of mature human adults, the synthesis of a manifold according to a schema *always* result in concept application? Why not think that manifolds are sometimes, perhaps often, schematised *without* the further step of subsumption under a concept being taken? As Longuenesse points out, not all representations are conscious for Kant and furthermore, not all representations we are conscious of are *thought* or amount to *experience*.

“...it is plausible to think that a great number of our representations involve acts of mere apprehension and reproductive associations of the same type as those animals are capable of, and the apprehension/reproduction/recognition *leading to concepts*- that is, to thought [experience] - is at work in only a small part of them”

(Longuenesse, 1998: 66)

Kant talks of representations which are not *representational* (Kitcher, 1990: 66; Hanna, 2001: 59). The natural way to interpret this claim is that just as in a developmental and comparative context we need to distinguish sensory apprehension undergirded solely by schemata from experience requiring also concepts, this distinction needs to be extending or applied to mature human adults also. That is, there may be sensory forms of awareness (perhaps in the sense of Travis, Ms.) in humans which are non-conceptual, but are merely schematic.

Insofar as Kant can admit this possibility,<sup>114</sup> he goes some way towards addressing the non-conceptualist intuition that there is a kind of sensory awareness which does not involve concepts. Of course, insofar as this sensory awareness is non-conceptual, *ipso facto* it cannot be properly called *perception* in the Kantian sense, for this is essentially experiential, objective. Nevertheless, this does not preclude it being called ‘perception’ in ordinary language; it may just be that ordinary language is insensitive to a technical difference between experiential perception, where manifolds are subsumed under concepts, and purely schematic sensory awareness. Despite his Conceptualism then, Kant’s distinction between concepts and schemata affords him considerable room for manoeuvre here. It enables him to admit of the possibility of a non-conceptual form of sensory awareness. His conceptualism is consistent with it.

Where the distinction between concepts and schemata has been neglected, the *Schematism* has inevitably been rejected as redundant. There is however, another less noticed, though equally inevitable, interpretive consequence of neglecting this distinction: concepts are ascribed the discriminatory role Kant attributes to schemata. Indeed, often this discriminatory role is taken to the primary function of concepts. For example Kukla (2011: 134) speaks of how a Kantian concept, “guides and regulates investigative behaviour and serves to inform the subject’s procedure *for determining whether the object of representation is a house or something else*” [my emphasis]. The picture, according to Kukla, is that what is received via the sensibility is consistent with being perceived in a number (perhaps an infinity) of different ways. The function of concepts is to privilege certain connections of sensible representations over others i.e. a concept serves to connect representations in X-way rather than Y-way. For example, Kukla (2011: 93) points out that the sensible representation of a house is often contiguous to the representation of the ground. Nevertheless, he says, we discriminate them perceptually and this is because we apply the concept of ‘house’, which does not extend to the ground. Kukla is by no means the only commentator to construe Kantian concepts in this way: as rules for connecting representations (e.g. Bird, 1962: 64;

---

<sup>114</sup> This is, surely, controversial. Kant is normally understood as claiming that all perception is essentially conceptual. On the view I am advancing by contrast, while concepts are certainly required for perceptual experience, i.e. that which exhibits intentionality in Kant’s sense, there is a distinct form of ‘sensory attunement to the world’ which is purely schematic and non-conceptual, and which, in ordinary language at least, would be called ‘perception’. It seems to me that this is a natural way to interpret Kant’s claim that not all representations need be or are brought to concepts. At the limit however, as I have already stressed, my claim is primarily that (a) Kant’s distinction of concepts and schemata provides him with the *resources* to take this view, and (b), given that this view deflects the objections to Conceptualism, Kant *should* take this view. Whether this can be truly regarded as a live exegetical option is ultimately a different matter.

McLear, 2011: 2). Pippin (1982: 91) describes concepts as the means by which we “apprehend and discriminate the world”, and Allison (2004: 187) suggests that Strawson adopts this view also. The problem is that these apprehension and discrimination locutions apply not to concepts, but, for Kant, to schemata. As Allison (2004: 210) stresses:

“it [a schema] functions to process the sensible data in a determinate way, thereby giving one a sense of what to look for or expect on the basis of certain perceptual clues.”

Crucially, this function of apprehension and discrimination is not one which requires *rules* or any reference to normative relations and so by identifying concepts with this function, the ‘rulishness’ of concepts inevitably disappears. This consequence is not obvious in Kukla (2011), since despite assigning concepts this discriminatory function, Kukla persists in describing concepts as ‘rules’. However, a close inspection of the function assigned to concepts shows that there is nothing essentially normative about them on Kukla’s account:

“the activity of perceiving the house is guided by a rule informing us of what we should or ought to detect in the course of scanning it...it is this temporally extended activity of perceiving, in the sense of *looking over*, that is unified by guidance from a rule” (2011: 96)

It is clear that what Kukla here labels a ‘rule’ is nothing other than a mechanical procedure: a series of steps which determine whether a manifold is X-way or not. Yet there is nothing normative about a mechanical procedure. Computing a function of this sort does not require normative language; no commitments need to be invoked. Therefore, despite Kukla’s repeated references to the normativity of concepts, if this is the role concepts perform, there is no good reason to construe them in normative terms at all; they are not *rules*.

The reason why this interpretive error, assigning concepts the function of schemata or equivalently, bleaching concepts of their normative aspect, is significant is that it renders concepts far less capable of performing the philosophical work Kant requires of them: securing the objectivity of experience.

“To subject such [sensory] states to rule-governed connections is to produce a *coherent and predictable system of ever-changing inner state*, which strikes me as required for [merely] representing them as states of the body” (Kukla, 2011: 129)

In attributing to concepts the role of schemata – roughly, that of bringing order to an indiscriminated spatio-temporal manifold - Kukla is unable to understand how the application of concepts achieves anything other than producing a “coherent and predictable system of the ever-changing inner state”. That is, he is unable to understand how the application of concepts is important in the way Kant claims: integral to bringing representations into relation with an object. In fact however, in misunderstanding Kant’s notion of a concept, Kukla fails to engage with Kant’s view at all. True, possessing a mechanical procedure for discriminating amongst one’s representations could not yield representation of an object, but since this is not what possession of a concept amounts to for Kant, Kukla is in no position to assess Kant’s claim that concepts are required for objective experience. It follows from this that distinguishing concepts from schemata is not only essential to overcoming the objections to Conceptualism, it is also inextricable from one of Kant’s most fundamental theses: that intentional experience requires concepts.

There is one more conclusion which warrants brief mention. Recall that I earlier agreed if the distinction between concepts and schemata is assessed solely as a response to the problem of the *Schematism*, the problem of concept application, then the distinction, arguably at least, collapses and the postulation of schemata as distinct theoretical objects is not justified. However, if the distinction is motivated on independently given grounds, as I have argued, it is much more reasonable to assume that schemata will be central to concept application. That is, insofar as the postulation of an independent sensory (or imaginative) capacity to discriminate amongst intuitions is required irrespective of any issue concerning concepts, it is reasonable to suppose that insofar as *there is* a problem regarding concept application, this capacity is likely to be part of the solution. Of course, this does not amount to an account of how this should proceed: a difficult matter still and one I shall not pursue here. Nevertheless, the suggestion is programmatic and makes better sense of the *Schematism*. In other words, the key to the *Schematism* is not to begin with the problem of concept application and then move to the postulation of schemata, but rather, to begin with the antecedently justified postulation of schemata and then move to the problem of concept application. Only in this case does the dialectic appear promising.

#### 5.4. *At the crossroads: meta-concepts*

I want to begin this section with a brief overview or summary of the philosophical thread which has been traced from the beginning of Chapter four to the present point. The original premise of this thread was that there is experience or objective cognition, understood in intentional and intensional terms. Adopting the transcendental method, the question became whether such experience presupposes a priori conditions for its possibility. It was observed that sensibility alone is insufficient for experience, and that the contribution of the understanding is required. The precise nature of this contribution was established to be that of concepts or, more to the point, rules. In turn it was argued that rules are only rules in the context of a judgement. This thread can be condensed into the following claims:

1. There is such a thing as intentional experience.
2. Intentional experience is necessarily rule-subjecting.
3. Rules require the capacity for judgement.

Now, in my view, there is one remaining claim to be added to this philosophical thread: one which concerns the existence of meta-concepts. To this point, with the exception of the distinction between event concepts and non-event concepts, I have spoken indiscriminately of concepts as rules, without drawing any distinctions or classifications within this class. The claim has been that experience presupposes concepts, but no particular concepts have been singled out as of special significance. Nevertheless, in this section I want to suggest the existence of just such a class of concepts: what I shall call ‘meta-concepts’.

According to Kant, judgements exhibit certain *forms*, which he organises into the follow table:<sup>115</sup>

---

<sup>115</sup> It may reasonably be wondered at this point where this table of forms comes from. What are the epistemic grounds for postulating this list of judgement forms rather than another, or indeed, none at all? Since my primary concern in this chapter is with the role Kant assigns to the forms of judgement, I shall not dwell upon the particular forms Kant postulates, or the method for their postulation. Instead, this latter topic is addressed in Chapter Seven.

Table 2.

<i>Quantity</i>	
<hr/>	
Universal	
Particular	
Singular	
<i>Quality</i>	<i>Relation</i>
<hr/>	<hr/>
Affirmative	Categorical
Negative	Hypothetical
Infinite	Disjunctive
<i>Modality</i>	
<hr/>	
Problematic	
Assertoric	
Apodeictic	

There are two issues regarding Kant's judgement forms which need to be addressed at this stage. Firstly, *what is* a form of judgement? What criterion need be satisfied for something to count as a judgement form? Macfarlane (2000: 90) is instructive in this respect:

“Kant thinks of this normative aspect of concepts and judgements – their rulishness – as their *form*.”

Just like concepts, judgement forms, for Kant, are rules. For instance, consider the following judgements:

- (1) ‘All men study Philosophy’
- (2) ‘One man studies Philosophy’
- (3) ‘One man does not study Philosophy’

For Kant, these judgements differ, in part, in virtue of realising different forms: whether they are universal or singular (1 vs. 2 & 3) and whether they are affirmative or negative (1 & 2 vs. 3). It is clear that these differences have direct implications for the normative or inferential profile of the respective judgements. For Kant then, the normative profile of any judgement is a function of (a) first-order concepts and (b) its form (A266/B322). In this sense, it makes perfect sense to think of these forms as rules: they serve to configure a subject's normative commitments.

Despite being rules however, Kant clearly means to distinguish judgement forms from other concepts. This brings us to the second point regarding the forms of judgement. According to Kant, every judgement realises one form from each of the four headings of his table: one form of quantity, of quality, of modality and of relation.<sup>116</sup> According to Kant then, there are a limited number of rules, some selection of which any judgement whatsoever must realise: and these rules he calls ‘forms’. More specifically, Kant believes not merely that judgements exhibit certain forms, but that there are a limited and complete number of forms which all judgements must exhibit (A79/B105). My present concern is not with the precise catalogue of these forms postulated by Kant, but rather with the role or function we ought to assign to such forms. That is, assuming Kant is correct in believing that there are a certain number of judgement forms, a certain number of rules which must be realised by any judgement, what is their philosophical significance?

I will address this question via a somewhat circuitous route. It has already been established that experience presupposes concepts. In turn, it has been established that concepts are only concepts (rules) in the context of a judgement. But now, suppose Kant is correct that all judgements presuppose certain forms. It would follow that the ability to apply these forms is presupposed by the use of *any* concepts. As such, we can label these forms ‘meta-concepts’, capturing both the rulishness of these forms, and their foundational relation to all other concepts. Furthermore, these meta-concepts cannot be known via experience since that would require the formation of a judgement which did not already invoke them: a posited impossibility. Consequently, these meta-concepts must be known a priori.

As a priori conditions of concept use however, it seems to follow that the philosophical significance of these meta-concepts is very great indeed. In fact, at this point, it seems that we are in a position to comprehend Kant’s astonishing claim that the intellectual conditions of experience can be derived, “almost by a single conclusion from the precisely determined definition of a judgement in general” (Meta: 145-6, fn. 3). Since all objects are grasped under some conceptual description, and all concept use presupposes the forms of judgement, these forms are intellectual conditions for the possibility of objective cognition: foundational for any relation to an object. To the extent that there are a limited number of judgement forms, the intellectual conditions of experience can indeed be derived from “the precisely determined definition of a judgement”.

---

<sup>116</sup> This is clearly something of an oversimplification since, for example, hypothetical judgements can have antecedents and consequences with different qualities and/or quantities.

“If we want to make them [representations] universally valid, [then] the representation is determined in *one or the other of the forms of judgement*; for example, I can myself think of motion and rest in succession, but if I say: “The motion follows on the rest,” then it is determined in the object and must be valid for everyone.” (Kant, *Metaphysik Volckmann*, in Guyer, 1987: 126) [my emphasis]

It is worth stressing that there is no threat of a regress here. If the argument was that grasping a concept necessarily presupposes the grasp of a foundational ‘meta-concept’, then a regress would indeed loom: in absence of some good reason to the contrary, it seems that the argument could be re-applied to the relevant meta-concept and so on, *ad infinitum*. However, this is not the argument. It is, I suppose, a conceptual possibility that there be no judgement forms, but that each judgement be merely a linear string, of some arbitrary length, of first-order concepts or rules: ‘green car fast’, ‘soon hit chair’ etc. The claim is not that there is some philosophical mistake involved in supposing the possibility of such formless judgements, it is instead simply that according to Kant, judgements *are not* formless, they always exhibit certain forms, and that insofar as concepts are only rules in the context of a judgement this validates the status of the relevant forms as meta-concepts.

Furthermore, not only is there no regress, it is possible that the forms of judgement are explanatory primitives, not admitting of any deeper account or rationale. While it is natural to ask why judgements exhibit the specific forms that they do, there might just be no answer to this. It might be that we can no more explain why judgement exhibits certain forms than we can explain why space and time are the forms of intuition (B145-6).<sup>117</sup> Insofar as we are willing to reconcile ourselves to this possibility, the account of experience which has been developed in this chapter would ‘bottom out’ with the forms of judgement: it would be as deep we can go. Equivalently, the existence of certain judgement forms would amount to the final point of Kant’s transcendental investigation: the culmination of the deduction beginning with the modest premise ‘there is such a thing as intentional experience’. As such, we could append to the claims:

1. There is such a thing as intentional experience.
2. Intentional experience is necessarily rule-subjecting.
3. Rules require the capacity for judgement.

---

<sup>117</sup> It is worth noting the parallels between this idea and Wolfram Hinzen’s suggestion that grammar constitutes an ‘anomaly’ in nature, in the sense that it is not reducible or explicable in non-grammatical terms. This comparison is significant since in Chapter Seven I argue that the table of the forms of judgement is grammatical.



, the following steps:

4. There are a limited number of judgement forms.
5. Judgement forms are meta-concepts.

Therefore: The forms of judgement are a priori conditions of experience.

This argument is more or less described by Stevenson (1979: 353):

“Walker informs us that as time went on, Kant got increasingly clear about how the argument should go. Not only does the notion of judgement come to the fore in B [the B edition of CPR], but ... in 1786 he had suggested that it could be carried out almost in a single step from the definition of judgement. So, Walker suggests, the essence of the argument is just this: all experience involves making judgements; the categories are the fundamental forms of judgement; therefore, all experience involves the categories.”

In my view, this is a plausible and attractive argument: one which, at the very least, ought to be taken seriously. Unfortunately - and this is a not a minor complication – it is one which Kant himself does not accept. To be precise, the final step of the argument is not one Kant takes. To continue the above quote from Stevenson:

“To make the second premise uncontroversially true [that the categories, the intellectual conditions for the possibility of objective cognition, are the fundamental forms of judgement], the categories must here be understood as... the *mere* logical concepts of negation, disjunction, and generality etc., not the more full-blooded... categories such as substance and causation.” [my emphasis]

Specifically, although Kant does ascribe the forms of judgement an important role, they are not, in and of themselves, considered to be of transcendental significance. In the next section I examine Kant’s conception of the forms of judgement and the function he ascribes to them in detail.

## Chapter 6. The Forms of Judgement

---

This chapter breaks with the foregoing discussion of Kant in a decisive respect. Up to this point, I have sought to outline a largely positive appraisal and reconstruction of Kant's views, forming something of a philosophical narrative, culminating in the summary of the previous section. By contrast, the present chapter is devoted to a critique of a central component of Kant's philosophy: his conception of the forms of judgement. I begin by locating Kant's conception of the forms of judgement in the context of the argument of the *Metaphysical Deduction*, since it is here that this conception is explicitly stated and philosophically invoked. The subsequent three sections are orientated towards illustrating the problems with this conception. I shall conclude that Kant's conception of the forms of judgement is mistaken, and that these forms ought to be instead understood in the terms outlined in the previous section (5.4). In a sense then, this chapter might be understood as an attempt to deal with a philosophical aberration of Kant's, and return us to the view of the forms of judgement already proposed.

### 6.1. The guiding thread of the *Metaphysical Deduction*

Although Strawson (1966: 82) famously pronounced of the *Metaphysical Deduction* that its results were not merely meagre, but so meagre as to render pointless detailed consideration of it, the argument has received renewed attention in recent years. Of particular importance has been Beatrice Longuenesse's *Kant and Capacity to Judge*, in which she argues that:

“[N]either the argument of the Transcendental Deduction of the Categories... nor the System of Principles of the Pure Understanding, can be understood unless they are related, down to the minutest details of their proofs to the role that Kant assigns [in the *Metaphysical Deduction*] to the logical forms of our judgements.” (Longuenesse, 1998: 9)

Kant begins the *Metaphysical Deduction* with the observation that judgements exhibit a limited number of forms and promptly codifies these forms into a table, as outlined in the previous section. Any judgement realises a certain selection of forms from this table. Now, the central claim or argument of the *Metaphysical Deduction* is that these judgement forms

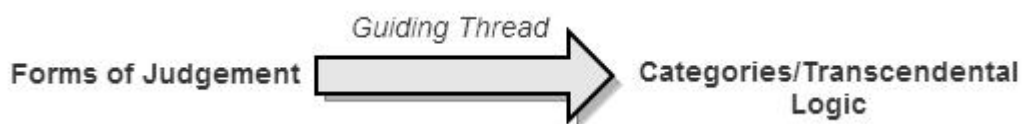
provide the ‘clue’ or ‘guiding thread’ to a *different* body of rules. The claim is that taking these forms of judgement as a premise, it is possible to derive or “systematically generate” (A80-81/B106) what Kant calls the *categories*.

“...there arise precisely the same number of pure concepts of the understanding... as, in the preceding table, there have been found to be logical functions in all possible judgements. For these functions specify the understanding completely, and yield an exhaustive inventory of its powers. These concepts we shall, with Aristotle, call *categories*.” (A79-80/B105)

‘Category’ is Kant’s term for a concept which applies to any object simply in virtue of being an object. Specifically, it denotes those rules to which objects must conform in order to be apprehended *as* objects: rules which must be applicable in order for cognition of objects or experience to be possible. Since such a concept cannot be founded upon experience, but, as a condition for the possibility of cognizing an object, is presupposed by experience, a category must be known a priori (A51/B75). In other words, categories are the intellectual conditions for the possibility of experience, corresponding to the sensible conditions, space and time, established in the *Transcendental Aesthetic*. Just as objects must be intuited spatially and temporally, categories denote whatever concepts objects must conform to in order to be *thought*. And just as Kant sometimes speaks of space and time not merely as the forms of any empirical intuition of an object, but as, themselves, constituting a pure intuition of an object (A20-1/B34-5), Kant similarly says that categories constitute a pure thought of an object (A55/B80). That is, categories delineate the contours of the pure thought of an object, absent any empirical component. Emphasising their transcendental significance, Kant labels the complete compendium of categories a ‘transcendental logic’: “a science of pure understanding... whereby we think objects completely a priori” (A57/B81). Transcendental logic comprises the rules according to which we must cognize objects. Given this, no cognition can contradict transcendental logic without losing all relation to an object (A62-3/B87).

As should be evident then, the categories occupy a position of paramount importance in the context of Kant’s broader transcendental undertaking. The claim of the *Metaphysical Deduction* is that the forms of judgement provide a clue to the categories: scrutinising these forms can unlock transcendental logic for us.

Figure 8.



There is a complication which lurks beneath this simple depiction. The *Metaphysical Deduction* contains no proof of the existence of categories: this comes later in the *Transcendental Deduction* and the *System of All Principles of Pure Understanding*. Hence, it is widely agreed by commentators that, notwithstanding Kant's occasional suggestion to the contrary, talk of the categories in the *Metaphysical Deduction* ought to be understood as 'anticipatory' (Guyer, 1987: 98-102).

"...hence we might well find it appropriate to distinguish also between pure and empirical thought of objects. *In that case* there would be a logic [i.e. transcendental logic] in which we would not abstract from all content of cognition." (A55/B79-80) [my emphasis]

Kant's talk of categories in the *Metaphysical Deduction* is 'anticipatory' insofar as he is making a claim about what must be the case if a certain possibility – that categories exist – turns out to be an actuality. That is, *if* there are such things as categories (a matter to be addressed elsewhere), then we are entitled to assume that they will correspond isomorphically to the forms of judgement (see Allison, 2004: 153). The claim is that although the existence of categories cannot be deduced simply by inspecting the forms of judgement, if there are categories then it is possible to enumerate them on the basis of the forms of judgement.

Even in this more qualified form, the central claim of the *Metaphysical Deduction* faces significant challenges. In precisely what sense do the forms of judgement provide the guiding thread to the categories? What justifies the derivation of one from the other? Clearly Kant considers there to be some connection here. Yet, it is not at all transparent what exactly this connection is and the matter receives little explicit attention in CPR. In the *Prolegomena* (§39, 71-72) Kant states that he simply "referred these functions of judging to objects in general... and there arose pure concepts of the understanding". It is therefore no surprise that the purported connection has been subject to considerable criticism (e.g. Bennett, 1966).<sup>118</sup>

<sup>118</sup> Moreover, it is unsurprising that commentators have defended incompatible accounts of how this derivation is supposed to proceed: compare, for example, Guyer (1987: 98) and Allison (2004: 156).

However, while the validity of the claim of the *Metaphysical Deduction*, - that the forms of judgement provide a clue to the categories - is a contentious and interesting matter, my concern here lies not with this claim per se, but with what it presupposes. Specifically, how must the forms of judgement be understood in order for the idea of a guiding thread or clue to be so much as coherent? Well, fairly trivially, this idea presupposes that the forms of judgement *are not themselves* categories: to say that the former can guide us to the latter entails that they are, in the first place, *distinct*, that they are disjoint domains.<sup>119</sup> And indeed, Kant's table of forms and table of categories are comprised of, apparently, quite different concepts. However, if the forms of judgement are not categories, not a priori concepts of an object, how should they be understood? The only alternative seems to be that they are a priori *non-objective* concepts. This is precisely Kant's view; they concern,

“...the form of discursive activity regardless of the object to which it may apply”  
(Longuenesse, 1998: 73)

This is reinforced by Kant's declaration that the forms of judgement constitute a *general logic*. In contrast to the transcendental logic the categories are said to comprise, a science of the pure thought of an object, Kant defines general logic in precisely dichotomous terms. It,

“abstracts from all content of cognition i.e., from all reference of cognition to its object. It examines only the logical form in the relation that cognitions have to one another, i.e. only the form of thought as such.” (A55/B79)

Kant's view is that bringing representations under the forms of judgement does nothing to secure a relation to an object and so, that these forms constitute only a general logic. While general logic contains “the absolutely necessary rules of thought without which the understanding cannot be used at all” (A52/B76), these rules concern only the *well-formedness* of a judgement, where being well-formed is one thing and relation to an object is another; it is only by applying a separate body of rules, the categories, that this relation is secured. At the heart of the *Metaphysical Deduction* then, is a fundamental dichotomy between, on the one hand, a priori non-objective or *formal* rules, the forms of judgement, which determine ‘well-formedness’, and on the other, a priori ‘object rules’, the categories, which establish a relation to an object. This dichotomy is integral to the argument of the *Metaphysical Deduction*. Not only does the notion of a ‘guiding thread’ presuppose a distinction of this kind, but also, it

---

<sup>119</sup> In other words, they are not intellectual conditions for the possibility of experience, as suggested in the previous section.

lies at the heart of what makes the argument provocative and interesting: namely, that a body of rules which are not, in and of themselves, of transcendental significance, the forms of judgement, can inform us of a body of rules which are of transcendental significance: the categories.

“...even while Kant makes extensive use of the “logical” forms of judgements... he also must distinguish categorical formality from what he calls merely “logical” formality” (Pippin, 1982: 90)

As such, the argument of the *Metaphysical Deduction* enshrines and fixes the view that the forms of judgement, in contrast to the categories, are not concerned with the relation of representations to an object, but are merely formal, non-objective, determining only the well-formedness of the judgement and nothing more. Hence, the argument presupposes that the conception of the forms of judgement proposed in the previous section (5.4), where these forms are understood to be, in and of themselves, of direct transcendental significance, is *false*.

My aim in the remainder of this chapter is to show that Kant’s conception of the forms of judgement as merely formal in this sense is mistaken. It is worth stressing from the outset then that despite the apparent terminological entailment, that the forms of judgement are *formal* is anything but tautologous for me. Rather I take it to be a substantive view: one which I shall be denying. I begin in the next section by discussing a traditional objection to Kant’s table of forms which appears unrelated to this issue. However, as I shall show, unravelling this objection leads inexorably to the question of the formality of the judgements forms.

## **6.2. Reappraising the forms of judgement**

A prominent objection to the argument of the *Metaphysical Deduction* has centred upon the purported ‘completeness’ of the table of forms from which the categories are to be derived. Kant emphasises that his table of forms is complete, describing it as “an exhaustive inventory of the powers of the understanding” (A79/B105). The claim is that each form is fundamental or primitive in the sense that they cannot be reduced to or replaced by another or some combination thereof. This is an important claim since, as Strawson (1966: 79) highlights:

“Given a certain indispensable minimum equipment of notions, the logician can, if he chooses, distinguish indefinitely many forms of propositions, all belonging to formal logic. If we allowed a category for each form, we should have indefinitely many categories.”

In other words, the table *must* contain only indispensable or primitive logical forms if it is to constitute the general logic from which the categories are to be derived. Yet this claim of indispensability has been subject to considerable criticism.

Firstly, many of the logical forms Kant includes in his table do not appear to be primitives. For example, Kant designates both ‘hypothetical’ and ‘disjunctive’ as distinct moments under the Relation heading. However, these forms are interdefinable, with the help of negation, in modern logic and so they cannot both be indispensable. Furthermore, Kant includes ‘infinite’ as a logical form, alongside ‘affirmative’ and ‘negative’ under the heading of Quality. This form concerns judgements such as ‘the soul is nonmortal’, i.e. where the predicate is negated. Yet this is clearly redundant, as Kant himself seems to admit when he says that it is only in transcendental logic, but not in general logic, that the infinite form must be distinguished from the affirmative (A73/B97). Yet, given that the thread of the *Metaphysical Deduction* is supposed to run from general logic to transcendental logic, to include the infinite form in the former on the basis of the latter is circular.

Secondly, setting aside reservations regarding the particular forms enunciated by Kant, there is a potentially far more damaging criticism of the table, one influentially pressed by Strawson. This is that the very idea of a complete table of logical forms involves a confusion: there could be no such thing. The reason for this is that:

“...as far as logical forms are concerned, the logician’s choice of primitives *is* a choice” (Strawson, 1966: 80)

In classical logic, there is no distinguished set of logical primitives because forms are inter-definable. For example, the same inferences can be modelled, with the help of negation, by utilising either disjunction or the conditional: neither is indispensable. Similarly, the existential quantifier and the universal quantifier can each be defined in terms of the other. Consequently, no system of logic – that is, a system of rules governing inferential relations

between truth-evaluable wholes – need take *any* particular form as primitive.<sup>120</sup> At most, Strawson argues, there may be certain primitive logical ideas: quantification and truth functional composition. Yet neither of these appears to be suitable for deriving any substantive categories for the pure thought of an object from. In sum, the objection to Kant’s table of forms is not simply that it is incomplete – many of the forms being dispensable –, but that the very idea of a complete and exhaustive inventory of logical forms is flawed.

The reason this completeness objection is important is that it has, in more recent years, prompted a reappraisal of Kant’s conception of judgement forms. The completeness objection assesses Kant’s table of forms in light of modern logic, and this evidently presupposes that the former is, in the relevant sense, intellectually continuous with the latter. If it is not then it is a mistake, at least in the absence of further argument, to critique the possibility of a complete table of forms from the perspective of contemporary logic. This is precisely what has been recently claimed by some commentators:

“...it must be insisted that the modern conception of logical form cannot be seen simply as a replacement for the Kantian conception and, therefore, cannot be appealed to in order to undermine the feasibility of Kant’s own project” (Allison, 2004: 146)

Kant’s enumeration of the table of judgement forms and contemporary logic do not constitute a single intellectual endeavour, but are separated by a “radical gulf” (Allison, 2004: 146). The claim is that while both concern the inferential properties of judgements, the notion of logical form which encodes for these properties is very different.

“Today we call “form” the structural features of the proposition that are relevant to truth-preserving inference and are expressed in the language of a logical calculus. In many ways Kant’s “formal” logic is even further from such a model than Aristotle’s *Prior Analytics*... His conception of logic is closer to that of the Port-Royal logicians, for whom logic was the exposition of “the reflections men have made on the four main operations of their mind-conceiving, judging, reasoning and ordering”.” (Longuenesse, 1998: 74)

In contemporary logic, assigning a logical form to a judgement is correct just in case it adequately captures the relevant inferential profile of the judgement. The task of the

---

<sup>120</sup> This point is not limited to classical logic. For example, in modal logic since necessity and possibility operators are inter-definable the logician can choose which to use (Cresswell, 2001).



philosophical logician is to devise systems which adequately encode for these inferential properties. Insofar as multiple logical forms can capture the same inferential profile of a judgement, neither is any more correct than the other: which is assigned is generally determined by meta-theoretical, pragmatic considerations (Quine, 1970). Accordingly, the logical form of a judgement does not purport to reflect any unique structure of the judgement itself, but is simply a regimentation or codification of certain inferential properties (cf. Borg, 2004: 62-73).

By contrast, commentators have argued that Kant conceives of logical form rather differently. Kant's table of judgement forms lists "forms of mental activity" (Longuenesse, 1998: 5) or "universal rules of discursive thinking" (Allison, 2004: 146) or a "depth grammar of thought and judgement" (Hanna, 2001: 79). The idea is that these forms are not simply regimentations imposed upon, or assigned to, judgements, but, in some sense, essential or intrinsic to judgements themselves. There is some fact of the matter as to whether a judgement has a particular logical form, even if there are other logical forms which are equally capable of capturing the inferential properties of the judgement, relative to some logical calculus. As such, Kant's notion of a judgement form is not continuous with that found in contemporary logic. In fact, it is arguably closer to the notion of logical form found in contemporary linguistic theory, where each sentence is taken to have a unique 'deep structure'.<sup>121</sup> Reappraising Kant's forms of judgement in this way serves to vitiate the completeness objection earlier described. Just because there could be no complete table of forms from the perspective of contemporary philosophical logic, this does not undermine the possibility that there could be a complete table of the 'depth grammar of judgement' (Hanna, 2001: 81).

In my view, this reappraisal of Kant's forms of judgement has significant implications, extending well beyond those relating to the completeness criticism. Specifically, it raises a problem for the thesis that the forms of judgement constitute a general logic. It opens up the possibility of their dissociation. Consider again how Kant describes general logic:

"As general logic, it... deals with nothing but the mere form of thinking." (A54/B78)

"General logic, as we have shown, abstracts from all content of cognition i.e., from all reference of cognition to its object." (A55/B79)

---

<sup>121</sup> Though nothing pivotal hangs upon this observation at this stage, the connection between Kant's forms of judgements and contemporary linguistic theory is developed in detail in detail in Chapter Seven.

Kant asserts that the forms of judgement constitute a general logic in this sense. That is, the forms of judgement are said to be a purely *formal* body of rules: they abstract from all relation to an object. Now, for the modern reader, this notion of a purely formal set of rules is perfectly familiar. Logic textbooks often open with the pronouncement that it is the form of an argument, rather than its content, which is at stake.

“Logic, whether classical or extra- or anti-classical, is concerned with form.” (Burgess, 2009: 2)

It is clear what notion of ‘form’ is relevant here. Modern logics stipulate rules for symbol manipulation, irrespective of what model theoretic interpretation is assigned to them. In this sense, any modern logic draws a principled distinction between form or syntax on the one hand, and content or relation to objects on the other. Given the success modern logicians have enjoyed in regimenting many of the inferential properties of language/thought according to such formal rules, it is natural that readers should find Kant’s declaration that the forms of judgement abstract from all relation to an object and so constitute a general logic unproblematic. That is, even if there are problems with the details of the table, the basic idea that the forms of judgement are ‘merely’ formal is perfectly acceptable and uncontroversial when viewed through the prism of contemporary philosophical logic. Indeed, of all the major elements of the *Metaphysical Deduction*, the formality of the judgement forms is perhaps the only one which has not been called into question. Worries have been raised with regards to the existence of transcendental logic, the relation between the table of judgement forms and transcendental logic, and, as has been discussed, the purported completeness of the table of forms. Yet, the *formality* or non-objective nature of the forms of judgement, that the forms of judgement constitute a general logic, - clearly a presupposition of the argument - is something which has never, to my knowledge, been critiqued. Rather this is generally accepted unquestioningly (e.g. Hanna, 2001: 78) and I am aware of no critical discussion of the issue.

However, a consequence of the reappraisal of Kant’s conception of judgement forms advanced by Allison and Longuenesse is that this uncritical acceptance is founded upon a mistake. Insofar as this reappraisal is sound then it is wrong to view Kant’s judgement forms through the prism of contemporary logic. Judgement forms are not simply regimentations of the inferential properties of a judgement, but stipulate “laws of the mind” or a “depth grammar of thought”. In *this* light however, the assertion that these laws are formal, non-

objective, is clearly not one which can be simply waved through. The scant recognition this point has received from commentators is perhaps partly attributable to the perceived absence of any alternative, or the belief that the assertion is essentially tautologous i.e. that it is a fact of language that judgement forms are formal. In the context of this thesis however, an alternative readily presents itself since I have already described it in Section 5.4: they are of transcendental significance. They are meta-concepts, applicable to all objects in virtue of being cognised as objects.

In sum, understanding Kant's table of forms through the prism of modern logic has shielded it from the question of whether these forms are formal; yet the reappraisal of Kant's conception of judgement forms renders this untenable. In other words, while the reappraisal may refute the completeness objection, stressing the 'radical gulf' between Kant's table of forms and contemporary philosophical logic comes at a cost. It entails that far from being something which can be nodded through uncritically, the assertion that the forms of judgement are formal, that they constitute a general logic, is one which needs justification and demands attention. As Macfarlane (2000: 56) argues, the claim that there are rules for thought *as such*, without any regard to objects, is a nontrivial claim, one which needs to be justified.

Interestingly, although this issue has been neglected by commentators, there is reason to think that Kant himself could not have been similarly oblivious to it. This is because, as Macfarlane (2000; 2002: 44-46) has persuasively argued, Kant's postulation of a general logic was an innovation and marked a significant break with how his predecessors had understood logic: "he self-consciousness adopts it, *against* the current of his time" (Macfarlane, 2000: 80). In particular, logic had not been conceived of as purely formal, but rather, as concerning, at least to a degree, the objective realm:

"I do not mean the logic of the [Scholastic] Schools... I mean instead the kind of logic which teaches us to direct our reason with a view to discovering the truths of which we are ignorant." (Descartes, 1644: 9)

"By logic or the art of reasoning I understand the art of using the understanding not only to judge proposed truth but to also discover hidden truth." (Leibniz, 1696: 463)

In other words, even if the modern reader is likely to accept uncritically the notion of purely formal rules, such acceptance was not something Kant could have been assured of. His attempt to secure this acceptance is a matter to which I now turn.

### 6.3. *Why Kant needs judgements of perception*

In the *Prolegomena*, Kant draws a distinction between judgements of perception and judgements of experience.

“Empirical judgements, so far as they have objective validity, are *judgements of experience*; but those which are only subjectively valid I name mere *judgements of perception*.” (Prol: §18, 45)

This distinction has long attracted interest because it appears to be inconsistent with the claims of the CPR (outlined in Section 5.2), in which judgement is defined specifically as:

“...a relation that is *objectively valid*, and so can be adequately distinguished from a relation of these same representations in which there would be only subjective validity” (B142)

This definition seems to flatly contradict the possibility of judgements which are only subjectively valid i.e. judgements of perception. Given that Kant does not mention perceptual judgements in the CPR, it has been common to dismiss the distinction as involving a terminological confusion (Bell, 2001: 7-8), or as being at odds with the arguments of the CPR (Guyer, 1987: 101). Even where the distinction has not been rejected however, it has seldom been ascribed any special significance or seen to do any substantive philosophical work for Kant. Defences have generally been limited to showing that, contrary to appearance, the distinction is *consistent* with Kant’s primary commitments in the CPR. For example, it has been common to subsume the distinction under the much broader critical dichotomy between subjective unity and objective unity (Bennet, 1966: 132-3), and attribute any residual surface tension between the definitions in the *Prolegomena* and the CPR to the differing aims and methodological constraints operative in the texts respectively (Allison, 2004: 178-182). At best though, such defences show that Kant is *able* to consistently draw the distinction, not that he ought to or need do so. By contrast, my contention is that the distinction between judgements of perception and of experience is crucial to Kant’s broader philosophical

position. Specifically, it is integral to the claim that the forms of judgements are purely formal or non-objective i.e. constitute a general logic, and as such, it is integral to the argument of the *Metaphysical Deduction*.

Consider first how Kant describes judgements of perception in the *Prolegomena*. They are said to be “valid simply for us, namely, for our subject” (§18, 45). That is, they do not demand that others agree with us or find things to be the same. Since, as has already been discussed, such universal necessity is the mark of objectivity for Kant, this entails that judgements of perception are *not objective*. Instead, they reflect only the perceptual state of the subject. “They only express a reference of two sensations to the same subject, namely, myself, and that only in my present state of perception, and are not therefore valid of objects” (Prol: §19, 46). In other words, judgements of perception do no more than reflect the ‘mere stream of impressions’ which experience is distinguished from and contrasted with. Instead, only judgements of experience are said to be objectively valid and, therefore, experiential.

The similarities between this division of judgements and the broader critical distinction between a merely subjective unity of impressions and an objective unity are striking. It is therefore not surprising that commentators have often viewed the judgemental division as a straightforward extension of the latter distinction. Nevertheless, it is critical to resist such an equation. The judgemental division is uniquely significant, over and above the more general theme of subjective versus objective unity because it entails the dissociation of objectivity from the forms of judgement. Critically, despite being ‘merely’ subjective, judgements of perception, like judgements of experience, are well-formed, in the sense of realising moments from the table of judgement forms (Longuenesse, 1998: 172). For instance, all of the examples Kant offers of judgements of perception are either categorical or hypothetical. As such, judgements of perception are marked by two properties: they are governed by the rules catalogued in the table of forms and they are subjectively valid: they lack any relation to an object. From this, we can derive an important conclusion:

“...although the logical forms of judgement can be employed in the former [judgements of perception], they do not suffice for the latter [judgements of experience]” (Guyer, 1987: 100)

But this seems to be exactly what Kant needs with regards to the argument of the *Metaphysical Deduction*. Specifically, it justifies the assumption which, in the previous section, I argued requires justification: that the forms of judgement are *formal*. Insofar as

judgements of perception are bound by the rules of the table of forms, and yet make no objective claim, those forms must fail to inform us of the objective realm. Such forms must be purely formal. In turn, this entails that in the case of judgements of experience, something of transcendental significance must be “superadded” (Prol: §18, 44) over and above the mere connection of representations according to forms found in judgements of perception; this ‘something’ must, of course, be the categories. In other words, the existence of judgements of perception entails that the forms of judgement constitute only a general logic, separate from transcendental logic proper.

This connection between the argument of the *Metaphysical Deduction* and judgements of perception has been almost universally neglected in the literature. Bird (1962: 106) comes close to recognising it when he notes that the postulation of judgements of perception shows that Kant would not have accepted the identification of judgement forms with the categories. But even here, there is no recognition on Bird’s part, or at least no expression, of how judgements of perception are *needed* in order to preclude this identification: that this identification looms as a threatening spectre over the *Metaphysical Deduction* and that Kant needs judgements of perception in order to dispel it.

In sum, I understand the matter as follows. If there are judgements of perception, then the assumption of the *Metaphysical Deduction* that the forms of judgement are formal is secure. No less significantly though, if there are *no* judgements of perception, this would drastically undermine the idea that the forms of judgement are purely formal. Very simply, if wherever there are forms of judgement, there is objectivity, i.e. if all judgements are judgements of experience, then there is reason to believe that the forms of judgement are not formal, but are of transcendental significance.

#### ***6.4. There are no judgements of perception***

There are two main questions which can be asked of judgements of perception:

- (a) Does Kant need judgements of perception?
- (b) Are there judgements of perception?

I have already given a strong affirmative answer to (a), and this evidently makes (b) all the more pressing.<sup>122</sup> This section is devoted to (b). The first part of the discussion will be limited to examining and assessing Kant's own examples of judgements of perception. Subsequently, the issue will be considered more generally.

Kant offers a number of examples of judgements of perception in the *Prolegomena*. Many of these examples can be grouped together as those involving the predication of a secondary quality: 'the room is warm', 'sugar is sweet', 'wormwood is unpleasant'. These judgements, says Kant, are "only subjectively valid" and so are merely perceptual judgements. The evidence Kant offers in favour of this view is that "I do not at all expect that I or any other person shall always find it as I now do" (Prol: §19, 46). Such evidence is stressed also by Longuenesse (1998: 172), who points out that we do not take such secondary quality judgements to "be valid at all times, for every empirical subject, in any circumstances". Kant's and Longuenesse's assertions here are surely correct: we do not consider the judgement 'the room is warm' to be valid at all times, for every empirical subject, in any circumstances. While being correct however, this observation is also completely irrelevant, and certainly does not constitute evidence for the judgements being perceptual. The reason for this is that no judgement, with the exception of a necessary truth, satisfies such a criterion. After all, to say that a judgement is valid in all circumstances is to say that there are no circumstances which could render it false i.e. it is a necessary truth. As such, if we take this criterion seriously, the distinction between judgements of perception and of experience collapses into the distinction of judgements which are necessarily true from those which are not. Since this is certainly not what Kant intended, we had better not take the criterion seriously.

If these secondary quality judgements are perceptual, as Kant claims, then they are no more than 'a reference of two sensations to the same subject': a reflection of a stream of impressions. Equivalently, if they are judgements of perception then they lack any reference to an object; they are bereft of objectivity. Helpfully, Kant's conception of objectivity provides us with a useful way to test this claim:

"Objective validity and necessary universal validity (for everyone) are exchangeable notions, and although we do not know the object in itself, yet when we regard a

---

<sup>122</sup> Unfortunately, (b) has been relatively neglected by commentators. This neglect is attributable, at least in part, to a failure to recognise the importance of judgements of perception to Kant's broader philosophical position; that is, it is largely because commentators have not answered (a) correctly that (b) has been neglected.

judgement as at once universal and necessary, objective validity is therewith understood.” (Prol: §19, 45)

For a judgement to be objective is equivalent to it being ‘universal’ or a candidate for general acceptance.<sup>123</sup> Given this, consider the following. Suppose that having walked into a room, I judge that the room is warm. Yet, only moments later, Julie who was sat in the room before my entry exclaims “this room is freezing!” Now, if ‘the room is warm’ were merely a reflection of sensations in me, Julie’s judgement should be utterly irrelevant to my own or, equivalently, my judgement should have no bearing upon her judgements. In fact however, Julie’s judgement is likely to prompt a range of behaviour on my part aimed at establishing why our judgements diverge. For example, I might check to see whether she is sitting next to a draughty window, or look to see whether I am adjacent to a radiator. Not only is searching for such explanations probable, it is, more importantly, *rational* in virtue of the two judgements we have made. Yet, were my judgement merely a reflection of sensation in me, making no claim beyond the incorrigibility of my own sensations, such behaviour would be rationally inexplicable. That this behaviour is not inexplicable, but rather, perfectly rational, shows that, ‘the room is warm’ is objectively valid: it is a *candidate for general acceptance* and the fact the Julie does not accept it warrants investigation of why she does not accept it. Encountering no radiator or draught, I might reply, “What do you mean? This room is warm!”

At this stage, Julie can respond in one of two ways. Firstly, Julie can renege upon her original position, perhaps answering, “Well, I *feel* cold anyways”,<sup>124</sup> in which case my judgement has been tacitly accepted. The second possible response is that Julie disputes my judgement. In this case I might show her the thermostat; point out that the temperature of the room is much higher than room temperature; remind her of an occasion when she was warm despite the temperature being lower than it is in the room. Although this would be an extraordinarily tedious dispute, it would be a dispute nonetheless: I am calling to her attention reasons to abandon her position and agree with my judgement; this clearly forecloses the idea that ‘the room is warm’ is merely subjectively valid.

---

<sup>123</sup> As should be clear this is a much weaker criterion than that of being ‘valid at all times, for all empirical subjects, in any circumstances’. Judgements can be candidates for general acceptance without satisfying this much more stringent criterion.

<sup>124</sup> It might be pointed out that it is possible that Julie is not reneging at all, but that by her original utterance “this room is freezing” she meant to express not the judgement ‘the room is freezing’, but a judgement like, ‘I feel cold, currently’. I set this possibility aside here since I address this issue more fully below.



There is another test which can be applied to determine whether secondary quality judgements are judgements of perception. Again taking as our point of departure the equivalency of ‘necessary universal validity’ and objective validity, we can ask whether Kant’s secondary quality judgements *commit* the subject to anything.<sup>125</sup> Suppose that I judge that sugar is sweet. This certainly seems to commit me to a whole range of beliefs. For example, it commits me to the belief that if a type of food contains a lot of sugar, then it will taste sweet.<sup>126</sup> It would be *wrong* for me to judge that sugar is sweet and judge that strawberries are very sugary, and yet deny that strawberries are sweet.

In sum, Kant’s secondary quality examples are not judgements of perception. Indeed, Kant himself appears to accept this later in *Logic* (§40, 119-120), where ‘the tower is red’ and ‘the stone is warm’ are both unequivocally categorised as judgements of *experience*. However, the other examples of judgements of perception in the *Prolegomena* appear, if anything, even less promising. One example is, ‘air is elastic’, where ‘elasticity’ refers to the property of matter to resist penetration (Longuenesse, 1998: 174). Kant assures us that this judgement can be perceptual: “I only refer the feelings in my senses to one another” (Prol: §19, 46). Yet this seems completely baseless, and the judgement would surely fail the two tests for perceptual status already discussed. The same applies to Kant’s final example of a judgement of perception in the *Prolegomena*: ‘if the sun shines on the stone, it grows warm’. This seems to exhibit necessary universal validity if anything does. Consequently, none of Kant’s examples of judgements of perception appear credible.

However, this is not the view Longuenesse takes of Kant’s examples, so it is important to consider her extensive discussion of the matter. Longuenesse argues that Kant’s examples ought to be understood in terms of, or assigned, certain ‘readings’:

- ‘The room is warm’ should be read as ‘it feels warm (pleasant) in the room’ (1998: 192).
- ‘Sugar is sweet’ should be read as ‘sugar tastes sweet’ (Ibid).
- ‘Air is elastic’ should be read as ‘air feels resistant to compression’ (1998: 174).

---

<sup>125</sup> A persistent theme in the previous chapter (esp. Section 5.1) was that these concepts: objectivity, necessity (in the normative sense of the term) and universality are all deeply entangled. In this light, I take it that this ‘normativity test’ is simply a different side of the same coin as the ‘universality test’ applied to the judgement, ‘the room is warm’.

<sup>126</sup> This is only a material commitment. It is conceivable, for example, that despite being sugary, strawberries also contain an acidic quality which prevents them from tasting sweet. The claim then is simply that ‘sugar is sweet’ and ‘strawberries are sugary’ provide me with a *pro tanto* reason to assent to ‘strawberries are sweet’.

- ‘If the sun shines on the stone, it grows warm’, should be read as ‘if the sun shining on the stone is among the objects I perceive at a given time, the stone getting warm is among the objects I perceive at a succeeding time’ (1998: 179).<sup>127</sup>

According to Longuenesse, these readings are all judgements of perception: destitute of objective validity. Although this could certainly be called into question, the more pressing worry at this point concerns the nature or status of the reformulations themselves. What does Longuenesse mean by the claim that Kant’s examples ‘should’ be read in these ways? Given that these readings are far from self-evident, what does this claim amount to? One possibility is that Longuenesse is simply eliminating Kant’s examples in favour of a different set altogether. However, it is clear that Longuenesse considers her observations to elaborate upon how Kant understood the matter. Her discussion is very much founded upon and continuous with Kant’s own examples.<sup>128</sup> In this case however, it seems quite mysterious as to what Longuenesse could intend by these reformulations. After all, ‘the room is warm’ is a different judgement from ‘it feels warm (pleasant) in the room’. Similarly, ‘if the sun shines on the stone, it grows warm’, is a different judgement from ‘if the sun shining on the stone is among the objects I perceive at a given time, the stone getting warm is among the objects I perceive at a succeeding time’. Insofar as these are different judgements altogether, how can they inform or establish the perceptual status of *Kant’s* examples? The key to this puzzle becomes visible when Longuenesse (1998: 174) says:

“I may suspend at least provisionally any claim to objective validity of such a judgement [air is elastic], and just *state* how things seem to me.” [my emphasis]

Significant here is reference to ‘*stating* how things seem’. This interest in the linguistic expression of a judgement pervades Longuenesse’s discussion of perceptual judgements (see also, 1998: 179, 192). The reformulations Longuenesse proposes are appropriately understood as containing a linguistic object on the left side of the ‘should be read as’, and a judgement on the right side. In other words, the point is that sometimes when we say ‘air is elastic’ we mean by this ‘air feels resistant to compression’. Similarly, in uttering ‘if the sun shines on the stone, it grows warm’, we *can express* the judgement ‘if the sun shining on the

<sup>127</sup> Longuenesse also uses the locution “may thus mean” (e.g. 1998: 174) in the stead of “should be read as”, in places. However, since these appear to be interchangeable for Longuenesse, I shall speak only of the latter for the sake of univocality.

<sup>128</sup> For example, Longuenesse (1998: 191-2) attempts to explain away seeming inconsistencies between the examples of judgements of perception Kant offers in the *Prolegomena* and those in *Logic*, something which would be inexplicable were her ‘readings’ supposed to amount to alternatives to Kant’s examples.

stone is among the objects I perceive at a given time, the stone getting warm is among the objects I perceive at a succeeding time'. As such, Longuenesse's central point is that the sentential correlates of Kant's examples can be used to express both judgements of experience and (what Longuenesse at least analyses as) perceptual judgements.

Supplementing this point, Longuenesse articulates an account of the circumstances which generally determine the type of judgement which such a sentence expresses. These circumstances concern the relevant epistemic grounds available to the subject at that point in time.

"Saying 'The stone is warm' because I now feel it to be warm is quite different from saying it after the complex process that may consist of (1) relating our judgement to other judgements on temperature of different objects in different circumstances, (2) comparing variations of temperatures constructed as intensive magnitudes."

(Longuenesse, 1998: 191)

The idea here is that the type of judgement we mean to express by an utterance is closely correlated with the epistemic grounds available to the subject.<sup>129</sup> For example, having just entered a room and feeling a wave of heat hit me, my utterance of 'the room is warm' is likely to express the perceptual judgement 'it feels warm in this room to me'. On the other hand, if I know that stoves produce heat and, upon entering the room, both feel warmth and see that the stove is lit, my utterance of 'the room is warm' is more likely to express the experiential judgement that the room is so and so (Longuenesse, 1998: 192). Similarly, Longuenesse says that if a subject's epistemic ground for uttering 'if the sun shines long

---

<sup>129</sup> In places, Longuenesse appears to take the stronger view that where a subject's epistemic position is relevantly restricted, judgements of experience are actually *impossible*. That is, where a subject's epistemic position is relevantly restricted, sentences of the form 'the room is warm' or 'air is elastic' can *only* express judgements of perception, rather than simply tending to so express: "Only after such a method [for augmenting one's epistemic state] has been systematically applied *can* a causal connection be asserted: 'the sun warms the stone' (Longuenesse, 1998: 179) [my emphasis]. There are glaring challenges for such a view. For one thing, it seems to suggest that expressing a judgement which one lacks epistemic warrant for is impossible. For example, assuming that the judgement 'all birds fly' is not warranted merely on the basis of seeing one bird flying, the view entails that such a judgement would be *impossible* – as opposed to simply unjustified – as a response to seeing a bird flying for the first time. Were one to say 'all birds fly' in such a situation, one *could not be* expressing the judgement 'all birds fly', but necessarily something of only subjective validity. At the very least, this is a surprising and counter-intuitive result. Whether this is Longuenesse's considered view is not clear to me – though her language sometimes suggests as much, she offers no arguments in favour of it. Either way, I set aside this question here since the view actually has no implications for the status of Kant's examples. Accepting that where a subject's epistemic position is relevantly restricted judgements of experience are impossible, is of no help in determining whether Kant's examples are judgements of experience or judgements of perception! It only entails that if they are judgements of experience, then they can be formed only once certain epistemic conditions are met, or that if they are judgements of perception then they can be formed without such epistemic conditions being met.

enough on a body, it grows warm' is only a limited number of a observed temporal successions holding of the sun shining on a body and the body growing warm, we are likely to mean only that 'if the sun shining on a body is among the objects I perceive at a given time, then the body getting warm is among the objects I perceive at a successive time'. In order to express the experiential correlate of this, it is necessary to "confront the correlations already obtained with many more, while perhaps also using the resources of mathematical constructions to anticipate and test further possible empirical correlations" (1998: 179). Longuenesse does not offer any concrete specification of 'how many' correlations would be sufficient in this respect. As such, her account is restricted to the modest point that the judgements we express are, to some unspecified degree, correlated with the relevant epistemic grounds available to a subject.

The problem with all of this is that it is incapable of establishing or, as far as I can see, even supporting the relevant claim: that Kant's examples are judgements of perception. While it is true that the utterance 'the room is warm' can be used to express the judgement 'the room feels warm', and while it is also true that such use is likely bound up with the epistemic state of the subject at the time, such facts are tangential to the question of whether Kant's examples are judgements of perception. All they establish is that certain utterances can express multiple judgements. This is a completely general phenomenon. If someone asks me whether I want a sandwich, I can utter 'I have just eaten' in order to convey or express the judgement that I do not want a sandwich. In other circumstances of course though, 'I have just eaten' just means I have just eaten! The extent to which utterances can express different judgements, the contextual factors which are relevant, and the distinctions which arise in this regard, is an interesting topic: one for philosophers of language (Borg, 2004). Yet insofar as Longuenesse's observations are simply a particular realisation of this general pragmatic fact, it is hard to see how they bear upon the purported perceptual status of Kant's examples. Longuenesse seems to think that because the utterance 'the room is warm' can express the apparently perceptual judgement 'it feels warm (pleasant) in this room' as well as Kant's judgement 'the room is warm', Kant's judgement must, in some sense, *be* a perceptual judgement also. But this is a non-sequitur and Longuenesse does not offer any argument in favour of it. Nor does Longuenesse's epistemic account alter this picture in any significant way. It may well be correct that certain utterances tend to express certain judgements depending upon the epistemic state of the subject. There is, however, nothing particularly novel or surprising about this since the epistemic state of the subject is part of the context

which fixes pragmatic interpretation of utterances generally. Longuenesse's account is of the form: in circumstances, C, an utterance, U, is more likely to express the judgement, J. No realisation of such a schema is capable of determining anything about the nature of the judgement itself.

None of Kant's examples of judgements of perception withstand critical scrutiny. Nevertheless, it may be wondered whether the argument against Kant's own examples, only serves to highlight that Kant chose his examples very badly indeed. Consider again some of Longuenesse's proposed 'readings' of Kant's examples: 'it feels warm (pleasant) in the room', 'if the sun shining on the stone is among the objects I perceive at a given time, the stone getting warm is among the objects I perceive at a succeeding time'. Rather than suppose that these 'readings' can tell us anything about Kant's examples, why not simply treat the readings themselves as examples of judgements of perception? After all, 'it feels warm (pleasant) in the room' and 'if the sun shining on the stone is among the objects I perceive at a given time, the stone getting warm is among the objects I perceive at a succeeding time' realise moments from Kant's table of forms and so are clearly viable candidates.

It is questionable whether even these examples are only subjectively valid in Kant's sense. Are Longuenesse's readings really no more than a mere reflection of impressions in the subject, utterly bereft of objective validity? It is far from clear to me that this is so, or whether they would pass the tests already discussed for perceptual status. However, pursuing this line of inquiry would rely upon fine margins, possibly too fine to be decisive to any interesting degree, so I set this aside here. Instead, my primary reason for doubting that these are judgements of perception lies elsewhere. The point of departure for this doubt is a straightforward puzzle: why did Kant not offer these examples himself? Even if the perceptual status of such judgements is open to dispute, what is surely indisputable is that these judgements are *better* candidates than any of Kant's examples. 'It feels warm (pleasant) in the room' is surely closer to being a mere reflection of sensations, being merely of subjective validity, than the judgement 'the room is warm'. Similarly, 'if the sun shining on the stone is among the objects I perceive at a given time, the stone getting warm is among the objects I perceive at a succeeding time', is a more credible example of a judgement of perception than Kant's, 'if the sun shines, it warms the stone'. This much is fairly obvious, but this obviousness makes Kant's choice of examples perplexing. This puzzle is exacerbated by the fact that there is nothing unusual or unfamiliar about Longuenesse's reformulations –

it is hardly credible that Kant simply failed to notice them. But then why did he ignore such examples, and proffer instead his own less plausible examples of judgements of perception?

In order to understand this decision, I think it is necessary to return to the philosophical purpose which judgements of perception are supposed to fulfil for Kant. Judgements of perception are needed to establish that the forms of judgement are formal: that being well formed does not ensure representation of an object.

“...the mere form of judgement is not sufficient to insure that the relation does hold of empirical objects” (Longuenesse, 1998: 178)

Yet, if judgements of perception are to establish this thesis, it is not sufficient to show that they are about states of the subject. As well as being subjectively valid, it is necessary that they be subjectively valid *merely in virtue* of being well formed. This sounds a little odd, but the idea is quite straightforward. It is clear that there are a whole range of conceptual devices which function so as to orientate a judgement upon the subject's own cognitive state rather than upon the world. For example, the phrasal adjuncts ‘it seems to me that...’ or ‘I have always found that...’ affixed to any judgement perform this kind of function. Similarly, the presence of some verbs seem to perform this function also, at least to a degree e.g. ‘taste’, ‘feel’. As such, were the question of the possibility of judgements of perception such that it could be answered in a positive fashion by examples of judgements containing such constituents, the question would be a fairly trivial one. Crucially however, judgements containing such constituents cannot count as evidence for the possibility of judgements of perception. The reason for this is that their existence proves nothing with respect to the formality of the forms of judgement. Consider, granting that the judgement, ‘it seems to me that flowers are pretty’ is subjectively valid, what does the possibility of such a judgement tell us about the formality of the forms of judgement? Well, nothing: the possibility of this judgement is perfectly consistent with denying that the forms of judgement are purely formal, or that they pertain only to some abstract notion of well-formedness which has no bearing upon the objectivity of cognition. This is because the subjective validity of the judgement can be plausibly attributed to the constituent concepts of the judgement, rather than being anything to do with the *form* of the judgement. Recall that for Kant, any judgement is comprised of (a) first-order concepts and (b) forms (A266/B322). The point then is that the subjective validity of these judgements can be plausibly attributed (a), therefore having no implications for the nature of (b). For example, the phrase, ‘it seems to me’ clearly plays this

role in the judgement ‘it seems to me that flowers are pretty’: if the phrase is eliminated, the subjective validity of the judgement disappears with it.<sup>130</sup> Yet insofar as the subjective validity of a judgement is attributable to such elements, the judgement has no implications for the nature of the forms of judgement. By contrast, what Kant needs to show is that *irrespective* of elements other than a judgement’s form, but simply in virtue of realising forms, simply in virtue of being well-formed, judgements are subjectively valid. This is what it means to say that judgements of perception must be subjectively valid merely in virtue of being well formed.

Now, I think that the reason why Kant did not proffer Longuenesse’s readings as examples of judgements of perception is that they fail this requirement. In my view all of Longuenesse’s readings, to whatever degree they are subjectively valid, pretty clearly depend upon the presence of specific conceptual constituents. I won’t reconsider all of the examples again, but only the one which, I think, is most plausibly construed as being a mere reflection of impressions in the subject:

‘If the sun shining on the stone is among the objects I perceive at a given time, the stone getting warm is among the objects I perceive at a succeeding time.’

Here, the relative clauses present in the antecedent and the consequent of the conditional are extraneous from the perspective of the judgement’s form: the judgement would realise the same forms from Kant’s table irrespective of their presence. Yet these constituents underpin the subjectivity of this judgement. If we drop them, we are left only with Kant’s example ‘if the sun is shining on the stone, the stone gets warm’, which I have already argued clearly exhibits necessary universal validity.

Consequently, Kant’s examples are not judgements of perception, and the most obvious examples of subjectively valid judgements do not qualify as judgements of perception. Although this does not constitute a complete survey of all the possible evidence, the striking paucity of clear examples of judgements of perception supports the conclusion that there are no such things. Now, in the previous section I argued that the existence of

---

<sup>130</sup> Determining whether a judgemental constituent functions in this way is a non-trivial methodological challenge. For adjuncts, the test is easy: see whether the subjective validity of the judgement perseveres in the absence of the phrase. The matter is trickier for concepts which are necessary for the well-formedness of the judgement since they cannot be dropped in this way, e.g. ‘feels’ in ‘the room feels warm’. In these cases, the test should be to substitute for the relevant concept, other concepts with an equivalent syntactic distribution. If the subjective validity of the judgement diminishes where such substitution is effected, it is clear that the subjective validity of the judgement must be a function of the original concept.

judgements of perception bears critically upon the issue of the formal status of judgement forms. So, the conclusion that there are no judgements of perception needs to be brought to bear upon this issue. Kant's view is that the forms of judgement leave us stranded with our own presentations: to bring representations under these forms moves us no closer to a relation with an object. Although realising such forms does constitute a certain well-formedness, this does not suffice to lay any claim to necessary universal validity. Instead, such validity emerges only where distinct rules, i.e. the categories, are 'superadded' to this well-formed construct.

"All of our judgements are at first [i.e. in the sense of being merely well-formed] mere judgements of perception... It is only subsequently that we give them a new reference, namely, to an object." (Prol: §18, 45)

However, the absence of any clear examples of judgements of perception strongly indicates that this account of judgement forms is deeply mistaken. Contrary to Kant's view, and specifically, contrary to the argument of the *Metaphysical Deduction*, the forms of judgement are not merely formal or non-objective. In fact, what we find is that wherever representations are brought under the forms of judgement, they are brought into relation with an object.<sup>131</sup>

"Every judgement claims a certain measure of objectivity within its self-chosen narrower sphere, no matter how limited its subject-concept. It is never satisfied with establishing a mere coexistence of representations, but it erects a functional coordination between them, so that whenever the one content is given the other is taken as required. The "is" of the copula is the expression of the connection."  
(Cassirer, 1953: 245-6)

---

<sup>131</sup> In my view, the reason why Longuenesse is so quick to uncritically accept Kant's examples as judgement of perception is that she never actually doubts the existence of judgement of perception, and the reason she never doubts their existence is that she unquestioningly accepts that the forms of judgement are purely formal.

"It may be objected that these judgements are in fact composed of judgements of experience... My view is that in order to understand Kant's analysis of the move from judgement of perception to judgement of experience, we must consider not the components, but the connective in the judgement ('If...then'). The question then becomes, How does the merely logical combination of perceptions expressed by this connective lead to the subsumption of intuition under the corresponding category."  
(Longuenesse, 1998: 176, fn. 20)

Longuenesse's view here can be summed up as follows: even if we struggle to identify examples of judgements of perception, this matters little since all we have to do is consider a judgement merely as a 'logical combination' of representations, as realising merely the forms of judgement. That is, *since* the forms of judgements are formal, a judgement merely adhering to these forms is necessarily a judgement of perception. From the present perspective of course, this simply begs the relevant question: namely, are the forms of judgement formal? In order to address this question (one Longuenesse does not seriously contemplate), we have to investigate the existence of judgements of perception rather than presume it.



The difference between representations which exhibit no objective validity, such as a mere list:

- {room, warmth}

or a complex idea:

- ‘the warmth of the room’, ‘the roominess of the warmth’

and representations which *do* relate to an object, is that the latter conform to or realise judgement forms:

- ‘The room is warm’, ‘if it is a room, then it is warm’, ‘either it is a room or is it warm’

This fact is utterly inconsistent with viewing the forms of judgement as formal or non-objective or securing merely some abstract notion of ‘well-formedness’ which has no bearing upon relation to an object. To the extent that judgements *do* approach subjective validity or, in Cassirer’s terms, claim objective validity within an especially ‘narrow sphere’, this is not a function of the form of the judgement, but of optional conceptual constituents of the judgement:

- (i) ‘Stones are warm’
- (ii) ‘It seems to me that stones are warm’
- (iii) ‘I have always found stones to be warm’
- (iv) ‘Whenever I have encountered the impression of stones in the past, the sensation of warmth has been contiguous’

Whereas Kant says that subjective validity is the default status of judgement *qua* realising judgement forms, *qua* being well formed, and that objective validity appears only with the superaddition of something else to this foundation, the reality is precisely the converse: moving from judgements (i) to (iv), we are trying to *overcome* the original objectivity undergirded by the judgement’s form. In other words, judgements, *qua* realising the forms of judgement, are objective.

This conclusion has two consequences. Firstly, this conclusion has an exegetical ramification insofar as it is capable of illuminating certain inconsistencies in Kant’s critical thought. For example, consider that despite foreclosing the possibility of judgements of perception, Cassirer, in the quote above, is nevertheless expressing a thoroughly Kantian idea.

In the CPR, Kant himself claims that it is the copula ‘is’, which distinguishes the objectivity of representations from the merely subjective (B142). I noted at the beginning of the previous section that Kant says things about the possibility of judgements of perception which are, if not plain inconsistent, at the very least, in great tension with each other. The reason I am reiterating this fact here is that we are now in a better position to explain this textual phenomenon. We can explain it not by dissolving the tension or explaining it away, but by agreeing that the textual tensions are real and substantial and explaining why Kant felt *unable* to adhere to a single uniform position. Specifically, he was unable, because he felt himself pulled in two different directions. I have made two principal claims regarding judgements of perception: (a) Kant needs them (in order to establish the formality of judgement forms), and (b) he cannot have them (there are no such things). One exegetical possibility is that while Kant recognised the importance of judgements of perception to his philosophical position, in certain periods he doubted whether they were available (as I have argued they are not). For example, the expulsion of judgements of perception from the *B Deduction* (despite being published four years after the *Prolegomena*) may indicate such doubt. Furthermore, although judgements of perception do reappear in *Logic*, the discussion is not obviously continuous with that found in the *Prolegomena*. As already noted, in the *Logic* Kant claims that examples such as ‘the stone is warm’ and ‘the tower is red’ are judgements of experience. Yet both of these are predications of secondary qualities, categorised in the *Prolegomena* as judgements of perception. This variance in examples may again indicate uncertainty on Kant’s part: that he realises that the security of judgement forms as formal rests upon the existence of judgements of perception and so he is ‘reaching’ for such judgements, but he is nevertheless unsure as to whether he has successfully provided any examples.

Secondly, and, in the context of this thesis at least, much more importantly, this conclusion amounts to the assertion (or re-assertion) of the conception of judgement forms proposed in Section 5.4. That is, the forms of judgement are, in their own right, meta-concepts: a priori, foundational rules for the thought of any object whatsoever. In this sense, despite according the forms of judgement a prominent role in CPR, Kant actually *understates* their importance. Far from being merely formal, non-objective rules, they are of inherent transcendental significance. In fact, their postulation should be understood as the final step of the argument thread beginning from the premise, ‘there is such a thing as experience’.

## Chapter 7. The Role of Grammar: Sketching a Research Program

---

In this final chapter, I discuss how Kant's table of forms can be updated in light of recent developments in Generative Grammar. The novelty of this suggestion is illustrated by Hinzen's (2012: 636) claim that:

“Introductions to the philosophy of language barely mention grammar as such, focusing upon words and their meanings instead.”

Given that grammar barely features in the philosophy of language, bringing it to bear upon Kant's theory of judgement forms is, at the very least, rather unexpected. My aim in this chapter is to rationalise the unexpected. Specifically, I argue that updating Kant's table of forms in light of developments in Generative Grammar is a viable and attractive undertaking. In other words, my aim in this chapter is to sketch a research program.

In Section 7.1, I begin by explaining the relevant notion of 'update', explaining what exactly it would be for Generative Grammar to be 'brought to bear' upon Kant's table of judgement forms. The subsequent sections, 7.2 and 7.3, concern the suitability of utilising developments in Generative Grammar in this way.

### ***7.1. Identifying the forms of judgement***

Recall that in Section 5.4, I outlined the following argument:

1. There is such a thing as intentional experience.
2. Intentional experience is necessarily rule-subjecting.
3. Rules require the capacity for judgement.
4. There are a limited number of judgement forms.
5. Judgement forms are meta-concepts.

Therefore: The forms of judgement are a priori conditions of experience.

In the previous chapter, I discussed how Kant denies (5), instead taking the forms of judgement to constitute a general logic. I argued that this denial is completely unjustified and that his alternative conception of the forms of judgement is not defensible. As such, I argued that Kant's eschewal of the above argument fails. Given this, we can treat Kant's failure to endorse this argument as an aberration, and return to it afresh.

At this stage however, a new issue flickers into focus: namely, how are the forms of judgement to be identified? (4) states that there are a limited number of such forms. However, this is only justifiable insofar as there is a secure and settled means for identifying them in the first place. So far, nothing has been said regarding these means. Unfortunately, despite its importance, Kant is almost entirely silent on this matter. In the CPR, Kant plucks the forms of judgement, seemingly, out of thin air. His table of forms is presented abruptly and without ceremony; we are then assured that this table is complete; and that is about it. Nor does Kant elaborate upon the matter in the *Prolegomena* where he states simply that with regards to the table of forms, "there lay already before me the entire, although not altogether faultless, work of the logicians" (Prol: §39, 71).

Despite Kant's appeal to the work of the logicians, it has been widely remarked that his table of forms was certainly not a re-capitulation of logics of the day (Hanna, 2001: 79), and as such, the spectre of arbitrariness hanging over his selection of forms is difficult to dispel. Longuenesse (1998: 74-78) argues that Kant's selection can be justified by recourse to the 'pattern of dependencies' which exist for Kant between general and transcendental logic.

"Kant asked himself which logical forms of judgement should be considered primitive *if the original* function of judgement is to "bring given cognitions to the objective unity of apperception," that is, to relate our representations to object." (Longuenesse, 1998: 78)

In my view, Longuenesse's argument cannot succeed since (a) it is unclear what justifies these dependencies and (b), such dependencies presuppose a clear demarcation of possible judgement forms which are then 'whittled down' once compared to the demands of transcendental logic. It is precisely this demarcation however, which is in question.

Ultimately, what is crucial from the current perspective is not the exegetical issue of whether Kant's catalogue of forms can be saved from the indictment of arbitrariness, but the philosophical one of how the forms of judgement are to be identified. Even if there is some

exegetical basis for re-assessing the apparent arbitrariness of Kant's table, it does not follow that Kant has a viable and plausible means for identifying the forms of judgement. Indeed, given that Kant has almost nothing to say about this philosophical issue, it seems reasonable to assume that we ought to be casting our net considerably wider.

Admittedly, Kant's appeal to the "work of the logicians" may inspire the idea that the contemporary philosophical logic is capable of identifying the forms of judgement. Nevertheless, considerations already outlined in Section 6.2 invalidate this idea. Specifically, whereas Kant's conception of judgement form is of something inherent to judgements, contemporary logicians view form as something imposed upon judgements: regimentations of certain aspects of the inferential profile of a judgement. Of course, what is considered part of a judgement's 'inferential profile' is itself a theory-relative matter. For example, the sense in which 'it is obligatory for John to attend school' warrants the judgement 'it is forbidden for John not to attend school' would be part of the inferential profile of the former according to a deontic logic, but not classical logic. The form ascribed to a judgement then depends upon the inferential profile the logician is interested in. In light of this however, to expect contemporary logicians to supply a 'depth grammar of thought and judgement', which is what Kant's forms of judgement are supposed to amount to, is not reasonable, and is only likely to produce confusion.<sup>132</sup>

Pivotal at this juncture is a comment made by Pippin (1982: 92). Pippin suggests that the forms of judgement are "what we would today call the forms of our language". This strikes me as a fascinating, programmatic suggestion: programmatic in the sense that it raises the possibility of identifying the forms of judgement by recourse to theories of *language*. Although this suggestion has not, to my knowledge, ever been developed or properly explored against the rich background of contemporary research on language, this is precisely what I shall attempt in this chapter. Specifically, I shall claim that Generative linguistics is well placed to supply the forms of judgement. The forms of judgement can be derived from

---

<sup>132</sup> Admittedly, it is a topic of debate in the philosophy of logic as to whether a principled distinction, i.e. one which is not theory-relative or relative to certain pragmatic considerations, can be drawn between content and form (cf. Macfarlane, 2000; for denial see: Whitely, 1951), perhaps the most plausible candidate for such a distinction being invariance under permutations of a domain. However, this debate is probably not as philosophically profound as it appears. The reason for this is that it is premised upon intuitions about which components of an argument ought to be ascribed to its form, and which components ascribed to its content in the first place: for instance, quantifiers on one side, and colour predicates on the other. These intuitions are the means by which principled distinctions are evaluated. Roughly, invariance under permutations of a domain is a 'good' candidate for distinguishing form from content because it maps onto our intuitions regarding what is form and what is content. Even if such a principled distinction is available, it is far from clear that this would compel us to view logical form as anything other than a regimentation imposed upon judgements.

our best theories of grammar. This is what I mean by updating Kant's table of forms in light of contemporary linguistic theory: our best theories of grammar can be the source of these forms, the means by which they are identified.

Pursuing an update of this kind is promising for a number of reasons. Firstly, and most obviously, it closes a rather glaring lacuna in Kant's position. Kant requires a means for identifying the forms of judgement, and this is exactly what the proposal promises. In this sense, it rescues the forms of judgement from a certain obscurity. Of course, Kant himself could not have intended or envisioned the forms of judgement being supplied in this manner, so the proposal clearly cannot be intended as a narrowly exegetical one. Still though, there is no reason why this proposal cannot feature in, and so strengthen, a philosophical reconstruction of Kant's position.

Secondly, this update promises to be a naturalisation of this component of Kant's philosophy. The study of grammar is a thoroughly naturalistic enterprise (Hinzen, 2006), so insofar as it is capable of identifying the forms of judgement, these forms are naturalistic: objects of naturalistic inquiry. Of course, the claim that at least parts of Kant's transcendental investigation can be understood in naturalistic terms is hardly a novel one. Kitcher (1990) famously embraces the label 'transcendental psychology' for her reading of Kant's position. Nevertheless, I want to stress one potentially significant difference between the kind of naturalisation I am proposing and that of Kitcher. One of the most prominent objections to reading Kant as a transcendental psychologist is that it excises from his account all traces of normativity. As Allison (2004: 147) says:

"...Kant's concern with mental acts is not to be construed in a psychological sense. Or, if one insists that *any* account of mental acts is by definition psychological, then the claim is that the account is not psychological in a pejorative sense."

This 'pejorative psychology', Allison continues, would be one which casts Kant as offering a "naturalized, empirical cognitive psychology, which undermines the essentially normative nature of his account of mental activity". As should be clear, the proposed naturalisation would not reduce Kant's position to that of 'pejorative psychology'. Far from excising normativity, the proposal is premised upon the thesis that experience is fundamentally normative or rule-subjecting in character. The proposal is that the forms of judgement or meta-concepts presupposed by any rule whatsoever can be uncovered by an investigation of grammar.

Thirdly, insofar as this update is successful, it has ramifications for how grammar is understood. In particular, I think that it is capable of informing understanding of *the role or function* of grammar. This suggestion may appear puzzling: surely the function of grammar is already well established? Indeed, in Section 3.2 I described grammar as a system for mapping or linking sound to meaning, that is, as a system of, or for, communication, so its function or role appears to be well determined.

However, while it is incontrovertible that communication is *a function* of grammar, whether this is its only function, or indeed, even its primary function is far from clear. Over the past decade, there has been increasing interest in the conjecture that grammar has a functional significance over and above that of merely mapping sound to meaning.<sup>133</sup>

“...language [grammar] is optimised for the system of thought, with mode of externalisation secondary” (Chomsky, 2007: 16-17)

The claim that grammar is somehow ‘optimised’ for thought, with ‘externalisation’, that is, the mapping from meaning to sound, being secondary is one which Chomsky has asserted repeatedly and despite the generality of this claim, it has direct implications for the functional significance of grammar. Specifically, it implies that the functional significance of grammar is not exhausted by the task of sound-meaning mapping, but is bound up with enabling the system of thought humans have available to them. As Chomsky (2005: 4) says, grammar is responsible for constructing the “infinite variety of internal structures that enter into thought, interpretation, and other human mental acts”. More precise formulations of this idea abound. Gomilla (2012) argues that grammar is responsible for the systematicity and productivity, understood in the Fodorean sense, of propositional thought. Fitch & Hauser (2004) suggest that grammar underlies the processing of phrase structure computations, and so serves to unlock the processing of, for example, the long distance dependencies which are unavailable to finite state computational systems. Hinzen and Sheehan (2013) argue that that grammar is a system of reference, such that any referential act is grammatical in nature.<sup>134</sup> Moreover, as well as being traced to ‘global’ processes such as reference and systematicity, grammar has also been hypothesised to underlie the processing of domain specific tasks. Katz and Pesetsky (Ms.) argue that the recognition of patterns of tonal tension and relaxation in music depends

<sup>133</sup> Although I certainly do not want to broach the issue here, it is worth noting that the renewal and sharpening of attention upon the functional significance of grammar has, for reasons which should be transparent, arisen alongside an explosion of interest in the question of language evolution (see Hauser, et al, 2002; Pinker and Jackendoff, 2005).

<sup>134</sup> This view is explored in some detail in Section 7.3.

upon subjects' grammatical competence (for a related proposal, see Mukherji, 2010, Chpts. 6 & 7). As well as musical cognition, it has been contended that grammar may yield arithmetical abilities (Chomsky, 2007a), and the act analysis integral to moral cognition (Kirkby and Reichard, Ms.).

The reason this taxonomy of hypothesised functions of grammar is significant is that it indicates that its role or function is far from being a settled or determined matter. Even setting aside, as I have done at this point, the credibility of these hypotheses, the mere fact that there is such debate regarding the matter creates the intellectual space for a novel hypothesis: that grammar functions so as to determine the forms of judgement, such that investigating grammar is the means by which the forms of judgement can be identified.<sup>135</sup> The reason I find this potential ramification or consequence of pursuing the proposed update so interesting is that it is a realisation of the kind of bi-directional relationship between philosophy and the sciences of the mind which originally motivated this thesis. Not only does the update use the empirical study of grammar to inform a philosophical issue, namely how the forms of judgement are to be identified, it also allows a broadly Kantian perspective regarding the forms of judgement to inform our understanding of the function of grammar. However, enumerating the philosophical consequences of this update or proposal is one thing. Determining whether it is plausible is another, and it is this latter matter which I now turn to.

## 7.2. *Linguistic variation*

One glaring question which must be addressed concerns linguistic variation. There can be little doubt that there exists considerable linguistic diversity and variation, and this appears to be immediately problematic for the present proposal. The proposal aims to update the forms of judgement in light of the forms of language. Yet, given linguistic diversity, there seems to be little reason to think that this would yield a single table of forms. Instead, such an update appears destined to yield a multitude of tables, perhaps one for each natural language, with no method for privileging one over the other. Indeed, I suspect the reason why Pippin's programmatic suggestion that Kant's judgement forms are what we would today consider to be forms of language has not been developed is due to the apparent variability of the latter.

---

<sup>135</sup> That the functional significance of grammar is far from being a settled matter, despite over half a century of empirical study yielding substantive results, offers a salutary lesson for how scientific inquiry proceeds (see Hinzen and Sheehan, 2013: 1, for a nice exposition of this point).



However, linguistic variation is an empirical matter, the precise nature and details of which cannot be pre-judged. Unfortunately, this fairly obvious directive has not always been heeded. For instance, Bennett (1966: 86) is comfortable dismissing from his armchair the idea that any language has means for expressing any particular kind of judgement. Although, he admits, language has so far been treated “glibly and superficially”, there is no reason to think that a more adequate account of it would support the thesis that languages universally encode for types of judgement. Ironically, this is exactly the kind of prejudicing of an empirical question which he goes on to rebuke Hume for indulging in (Bennett, 1966: 96). What is required is an inspection of the nature of linguistic variation; only then is it possible to determine whether it invalidates updating the forms of judgement by recourse to the forms of language.

Prior to the inception of Generative Grammar, it was widely held that there are no bounds upon the variation natural languages can exhibit. Joos (1966: 228) famously declared that “languages differ from each other without limit and in unpredictable ways”. However, one of the most formative and important pillars of the generative tradition has been that this assumption of unbounded variation is empirically false. It has instead been observed that natural language variation is non-trivially constrained. For example, all languages are hierarchically rather than linearly structured, exhibit embedding and constituent displacement,<sup>136</sup> and admit of fundamental distinctions such as that of heads and non-heads, arguments and adjuncts. Furthermore, there are implicational relations between distinct typological phenomena e.g. the ordering of verbs and objects, and the ordering of adpositions and their complements (Greenberg, 1963). Such limits upon linguistic diversity render the topic of linguistic variation amenable to theoretical investigation. Historically, explaining or accounting for the constrained variation exhibited by languages has been a central challenge for generative linguists (Rizzi, 1982: 117): why do languages vary in the ways that they do?

The most influential response to this challenge within Generative Grammar has been that of Principles and Parameters (P&P). This model postulates that language acquisition is underpinned and guided by a genetic endowment for language: a Universal Grammar. Crucially, according P&P, this Universal Grammar consists of a number of parameters which

---

<sup>136</sup> Though, see Everett (2005).

are valued as part of the acquisition process.<sup>137</sup> Consider, for example, the null subject parameter:

‘Allow/do not allow finite verbs to have null subjects’

On the P&P model, acquisition of a grammatical competence is a matter of valuing a relatively small number, perhaps between thirty and forty, of such innate parameters. Moreover, each parameter is supposed to be binary, in the sense that only two possible values attach to each parameter (Newmeyer, 2005). This model is captured by the famous metaphor of a switchboard: to acquire a grammar is to set a pre-determined number of switches on a board.

“What we expect to find then, is a highly structured theory of UG based on a number of fundamental principles... with parameters that have to be fixed by experience.”  
(Chomsky, 1981: 3-4)

A consequence of this model is that the learning task faced by children in acquiring a grammar is simplified considerably. As such, the model has been widely viewed as a viable solution to ‘Plato’s problem’ or the problem of the poverty of the stimulus in the linguistic domain.<sup>138</sup> Furthermore, and, from the present perspective, more importantly, this model also serves to account for and explain a wide range of linguistic variation. Quite straightforwardly, the space of possible language variation is determined by the range of parameter values made available by Universal Grammar. Moreover, it has been hypothesised that parameters are hierarchically organised, such that the valuation of one has direct consequences for the valuation of others (Baker, 2001). Such ‘cascade effects’ (Chomsky, 1981: 6) could account for implicational relations found to hold of distinct typological phenomena (as well as predict previous unnoticed typological clusterings).

The influence of this Principles and Parameters model has been such that it has been applied *mutatis mutandis* to variation found in other domains e.g. moral variation (Harman, 2008). However, despite its influence, there has been growing recognition that, with respect

---

<sup>137</sup> Such parameters must be distinguished from the ‘micro-parameters’ much discussed in recent literature, and which attach to lexical items.

<sup>138</sup> This problem has been much discussed in the literature, and I will not dwell upon it here (see Pinker, 1994 for an overview). In brief, the linguistic data children have available to them is incomplete and awash with sentence fragments, idiomatic usages, and performance errors or slips of the tongue (cf. Hinzen, 2006: 122-124). Yet, in normal environmental conditions, children progress linguistically very rapidly, systematically avoiding the kind of over-generalisations which would be expected of domain general learning (Crain and Pietroski, 2005).

to language at least, the P&P model faces serious, perhaps insuperable, difficulties. Aside from a core cluster of parameters, there has been little progress in determining an adequate taxonomy of parameters (Haspelmath, 2008). On the other hand, it has been argued that there is little reason to think that relevant variation can be accounted for by a small number of parameters or that merely binary parameters can suffice (Newmeyer, 2005). For example, Manzini and Wexler (1987: 419) contend that cross-linguistic variation in pronominal and anaphoric binding can be accounted for only by a Governing Category parameter which can be valued in one of five distinct ways:

$\gamma$  is a governing category for  $\alpha$  iff

$\gamma$  is the minimal category that contains  $\alpha$  and a governor for  $\alpha$  and has:

- a. a subject; or
- b. an infl.; or
- c. a tense; or
- d. a “referential tense” tense; or
- e. a “root” tense

From the present perspective, what is significant is not criticism of the Principles and Parameters model per se, but with how this criticism has coincided with a radical re-appraisal of the challenge this model was supposed to address: that of explaining linguistic variation.<sup>139</sup> Traditionally, linguistic variation has been understood as, at least in part, a grammatical phenomenon, and as such, has been assumed to be a challenge for the grammarian, to be explained by theories of grammar. For instance, insofar as the P&P model seeks to explain linguistic variation in terms of differences in grammar parameterisation, it is evidently premised upon and enshrines this assumption. As Hinzen and Sheehan (2013: 301) put it:

“Grammar... is what has been the primary focus of the field of comparative syntax in the Principles and Parameters (P&P) tradition.”

However, in recent years this assumption has been increasingly called into question. It has been argued that rather than treating linguistic variation as a grammatical phenomenon, and so a problem or challenge for grammarians, such variation is best analysed in terms of differences in *externalisation*. ‘Externalisation’ broadly refers to the morpho-phonological

<sup>139</sup> I do not mean to imply that problems with P&P solely precipitated this shift. Although there is, I think, a connection, Boeckx (2011) argues that Minimalism, which has dominated the study of language for the past two decades, necessitates a reappraisal of the notion of linguistic variation (specifically, it implies that narrow syntax does not vary) on independent grounds.

realisation of grammatical properties. It has long been recognised that the morpho-phonological realisation of grammatical properties varies across languages. For example, it is widely agreed that nominals in English are assigned structural case. Yet, unlike many languages, with the exception of first and third person personal pronouns, English does not mark for case either lexically or morphologically. As such, there is a dissociation of grammar and its externalisation. The basic thought in relation to linguistic variation then, is that such variation can be traced to how grammar is externalised, rather than grammar itself.

For example, head-complement order and null subjects, two prominent points of linguistic variation once attributed to differences in parametric valuation, have been reassessed as different externalisations of a common underlying grammar. Languages vary as to whether heads precede or follow their complement. For example, in English heads precede their complements:

[read [the book]]

[in [the room]]

, whereas in Japanese they follow them:

[[book] read]

[[room] in]

Although in the P&P model this variation was explained by citing differences in how grammars are parameterised i.e. X precedes/follows its complement, the consensus view nowadays is that linearization or word order is not a grammatical matter at all, but wholly restricted to how grammatical hierarchies are externalised (Berwick and Chomsky, 2011). As such, the relevant variation is attributable to differences in phonological spell out. Null subject variation concerns whether languages require overt subjects in finite clauses: for instance, Chinese and Italian do not require such subjects, English does. Although this variation has been analysed in terms of a ‘null subject parameter’, Hinzen and Sheehan (2013: Chpt. 5) argue that, as with head order, the existence of null subjects is not a grammatical phenomenon at all, and that such variation is best understood as differences in externalisation.

Such developments raise the distinct possibility that all linguistic variation can be attributed to externalisation: word order, morphological operations and vocabulary insertion (Berwick and Chomsky, 2011; Cinque, 1999: 141). In turn, ascribing linguistic variation to

the non-grammatical components of language renders it highly plausible that grammar itself admits of *no* variation. It is now often asserted that, at their grammatical core, all languages are uniform (Chomsky, 2007a; Hinzen and Sheehan, 2013: Chpt. 5).

“...there is only one language, Human” (Boeckx, 2011: 210)

Returning to the original concern of this section, how can deriving the forms of judgement from the forms of language be possible given linguistic variation, two points can be extracted from the foregoing discussion. Firstly, the notion of ‘linguistic variation’ is insufficiently fine-grained for our purposes because it elides the difference between variation attributable to externalisation and that attributable to grammar. This matters because the forms of language targeted by the proposed update are those of grammar, those uncovered by generative linguists, and these must be distinguished from their morphological and phonological realisation. ‘Forms of language’ here means forms of *grammar*.<sup>140</sup> As such, it is not linguistic variation *tout court* which threatens the proposed update, but grammatical variation in particular. The second point which follows from the foregoing discussion is that there may well be no such thing as grammatical variation. Consequently, the objection, properly formulated, has been nullified.

### 7.3. *Un-Cartesian linguistics*

Even if grammar is uniform, there is another major obstacle for the proposal that it supplies the forms of judgement. This obstacle concerns the supposed autonomy and arbitrariness of grammar. Since the inception of Generative Grammar in the 1950s, the orthodox conception of grammar has been that of an *autonomous domain*: comprised of rules and principles that can be studied in relative abstraction from cognition at large. Grammar has been understood according to the metaphor of a ‘mental organ’, a language faculty or module in a roughly Fodorean sense (Pinker, 1994). As a module, grammatical processing is isolated from and independent of that executed by other modules and also, crucially, independent of any non-modular understanding: plausibly, cognition or thought, that which bears semantic properties (Fodor, 1983). In this sense, grammar constitutes an autonomous domain of inquiry. As Hinzen and Sheehan (2013: 404) put it, Generative linguists take themselves to be studying,

---

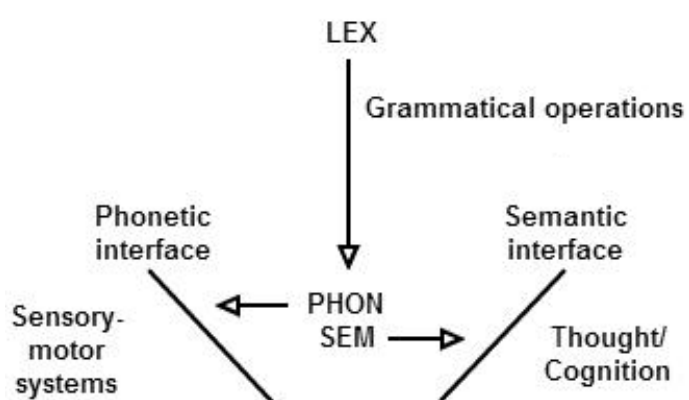
<sup>140</sup> So, although I shall continue to speak of ‘forms of language’, this ought to be understood as referring to the forms of grammar.

“not thought, but an autonomous structure that can (or must) be studied independently from any general ‘cognition’”.

The notion of arbitrariness central to this orthodox conception of grammar is that of *semantic* arbitrariness. Arbitrariness of this kind follows from the autonomy of grammar. Essentially, if grammar is one thing and that which bears semantic properties, thought or cognition, is another, then grammaticality can be of no semantic significance: in this sense, grammar is arbitrary.

This orthodoxy of autonomy and arbitrariness is closely connected to the role of grammar in mapping sound to meaning, and is perhaps best illustrated in light of how this mapping function has been modelled. According to linguists, grammar maps sound to meaning insofar as it generates a *pairing* of phonological information and semantic information:

Figure 9.



Essentially, any grammatical derivation begins with, or takes as its input, a set of unordered words, bundles of phonological, syntactic and semantic features, from the subjects’ mental lexicon: LEX. Grammar combines these words as a function of (a) their syntactic features, and (b) the rules of the system. Upon completion of the derivation, that is, if all the relevant grammatical rules have been successfully met, the result is a pairing of phonological information, PHON, and semantic information, SEM. These two kinds of information are then transferred or made accessible to ‘external’ cognitive systems responsible for articulating sound and those of ‘thought’ respectively.<sup>141</sup>

<sup>141</sup> This model oversimplifies matters slightly. In particular, it omits the fact that transfer to the interface occurs at the phasal, rather than the derivational level (cf. Hinzen, 2012a)

This model enshrines the autonomy and semantic arbitrariness of grammar. In explaining how grammar maps sound to meaning, linguists postulate a distinct system(s) responsible for meaning, thought/cognition, and a distinct system(s) for sound, sensory motor system(s). Grammar simply relates or mediates between these two antecedently postulated systems, systems which exist independently of being so mediated. In explaining how grammar maps sound to meaning then, linguists distinguish the mapping *from that which is mapped*, and as a consequence, guarantee the autonomy of grammar. The semantic arbitrariness of grammar is equally clear on this model, where meaning is ascribed to an ‘external’ system of thought with which grammar merely interfaces. Far from grammar being of any inherent semantic significance, in mapping sound and meaning, grammar is taken to presuppose the independent availability of a domain of meaning such that it can be mapped at all. The rules according to which grammar pairs sound and meaning are rich indeed, yet they have no implications for the nature of that which is being paired. For linguists then, grammar is a non-semantic domain; the rules of grammar are semantically arbitrary.

Indeed, the semantic arbitrariness of grammar was precisely what Chomsky’s (1957) famous example, ‘colourless green ideas sleep furiously’, purported to show. Such a sentence, says Chomsky, is “nonsense”, in the sense that it is not a coherent or meaningful thought. Nevertheless, it is absolutely grammatical for speakers of English. What this indicates is that the rules of grammar have no bearing upon what counts as a meaningful thought.

The reason this is significant from the present perspective, is that accepting the autonomy and arbitrariness of grammar renders it incapable of supplying the forms of judgement. To say that grammatical operations are semantically arbitrary is to say, from a Kantian perspective, that they have no bearing upon the relation of representations to an object. As such, as long as grammar is understood to be semantically arbitrary, the proposal that the forms of judgement can be updated in light of our best theories of grammar is not at all plausible.

Now, it is certainly true that this ‘orthodoxy’ is increasingly less orthodox, at least in a key respect. In Section 7.1, I outlined how it is increasingly common for the mapping from meaning to sound to be viewed as only *a* function of grammar, and perhaps a secondary one. The alternate functions postulated coalesce around the idea that grammar enhances the kind of computational processing subjects are capable of in various domains. Specifically, it is said to yield recursive computation, that is, the embedding of a constituent in a constituent of

the same category, in other domains. In current minimalist linguistics, the core operation of grammar is that of Merge. Merge takes two objects, X and Y, and forms the unordered set, {X, Y}. This set may be thus merged again with a further object Z, whereby the structure {Z, {X, Y}} is created.<sup>142</sup> Merge is a simple operation. Yet, with simplicity comes computational power: trivially, merge is a recursive operation. As a consequence, where recursion has been found in putatively non-grammatical domains such as the language of thought (Berwick and Chomsky, 2011; Gomilla, 2012), arithmetic (Chomsky, 2007a), moral cognition (Kirkby and Reichard, Ms) and musical cognition (Katz and Pesetsky, Ms), this recursive computation has been explained by recourse to subjects possession of a grammatical competence.

A few observations are in order at this point. Firstly, these claims amount to repudiations of the autonomy of grammar. Insofar as grammar has consequences for a wide range of cognitive domains, the study of grammar is not independent of or distinct from study of these domains. For example, if Kirkby and Reichard (Ms.) are correct, moral psychologists are studying, in part, grammar. Secondly, it is not surprising that, *to the extent that* grammar has been claimed to be non-autonomous, its significance has been traced to its computational ramifications. Merge aside, the computational paradigm is central to much contemporary theorising about grammar. Grammatical derivations are taken to be driven by the requirement that certain formal features attaching to lexical items, such as case and  $\phi$  features (number, person, gender), be ‘checked’, where checking is a thoroughly computational challenge, bound up with such concepts as ‘search’, ‘delete’ and ‘crash’. Similarly, computational efficiency is widely invoked as an explanation of various grammatical phenomena, such as, for instance, how movement only seems occur as a ‘last resort’ (cf. Hinzen, 2012a). In other words, the grammatical paradigm is a computational one, and so it is hardly surprising that the hypothesised effects of possessing a grammar have converged upon the kind of computations a subject can process.

However, while the claim that grammar underpins computational processing in distinct domains does vitiate a very strong modular or autonomous conception of grammar, it does little to enhance the plausibility of the proposal that the forms of judgement can be updated in light of grammatical theory. The reason for this is that the computational paradigm is founded upon a categorical distinction of semantics and syntax. Computational processes

---

<sup>142</sup> I am here eliding an issue which is, strictly speaking, crucial to the recursive nature of Merge: that of labelling or headedness. Since this technical point is not of any significance to the present discussion, I shall not explore this matter here (cf. Hinzen, 2006: 170-94).



are sensitive to the physical properties of symbols, not what they denote. Grammar may yield recursive computation in a number of domains, but such an effect is necessarily syntactic, and is of no inherent semantic significance. In other words, grammar is still semantically arbitrary. Yet it is precisely semantic properties which matter from the perspective of the relation of representations to an object. As such, if the effects of grammar are ‘merely’ syntactic, there is little hope for updating the forms of judgement in light of linguistic theory.

Crucially however, there is a striking and increasingly well developed, alternate conception of grammar, one which breaks with autonomy in a much more radical sense: the Un-Cartesian view (Hinzen, 2009, 2012; Hinzen and Sheehan, 2013; Sheehan and Hinzen, 2012).<sup>143</sup> This view denies the distinction of grammar and thought/cognition altogether. Instead, the study of grammar *is* the study of thought, and *a fortiori* there can be no “interface” between them (Hinzen and Sheehan, 2013: 417).<sup>144</sup> On this view, it is not simply that grammar is non-autonomous in the sense of “constraining” or having computational consequences for cognition (Berwick and Chomsky, 2011). Rather, grammar is non-autonomous in the sense that there is no independent system of thought for grammar to constrain. Thought is inherently grammatical.

Even this cursory description of the Un-Cartesian position is sufficient to convey its radical character. It has long been widely urged that language and thought must be distinct. Indeed, the idea that thought might be, in some sense, distinctively linguistic or grammatical has been commonly viewed as not merely false, but dangerously confused (Fodor, 2001).

“...the idea that thought is the same as language is an example of what can be called a conventional absurdity” (Pinker, 1994: 57)

Jackendoff (1996) summarises the reasons why language and thought must be separate. Perhaps the most widely cited of these reasons is that thought is immune to differences in language. For instance, Jackendoff notes how we do not suppose that the fact that someone speaks French means that she has different kinds of thoughts from that of a Turkish speaker. Rather, thought is independent of the language spoken, as demonstrated by the possibility of translation. Translation presupposes that different languages can express the same thought. Yet, were it the case that a French speaker not merely spoke in French, but thought in French,

<sup>143</sup> For a related proposal, the Representational Hypothesis, see Burton-Roberts, 2011.

<sup>144</sup> It is worth noting that this thesis renders the minimalist strategy of rationalising language design by recourse to the ‘legibility conditions’ imposed upon the grammatical system by separate conceptual-intentional system(s) largely incoherent (e.g. Chomsky, 2005).

*mutatis mutandis*, for other languages, the idea of translation would be incoherent, so goes the argument.

However, from the present perspective, this argument misses its mark. As was stressed in the previous section, the relevant conception of language or grammar is one which admits of no variation. While social constructs such as French and Turkish do indeed differ, grammar is defined at a much higher level of abstraction, one which *ex hypothesi* is completely uniform. Consequently, language variation cannot be used to support the separation of grammar and thought.

Another consideration which has been marshalled in support of the separation of language and thought concerns comparative cognition. Animal communication systems are bereft of grammatical structure (Cheney and Seyfarth, 1997). Furthermore, attempts to teach apes (including chimpanzees) to sign human languages have foundered most strikingly with respect to grammatical relations (Terrace, 2005).<sup>145</sup> Grasping grammatical relations appears to lie beyond the reach of animals. Despite lacking grammatical competence however, so goes the argument, animals think. For instance, Jackendoff (1996) points to the ability of vervet monkeys to ‘reason’ about appropriate retribution. Roughly, if monkey X attacks monkey Y, there is a strong likelihood of Y thereafter attacking some member of X’s kinship group: a fact which, Jackendoff suggests, implicates some fairly sophisticated reasoning on the part the monkey about what constitutes appropriate retribution. Ergo, grammar and thought must be distinct.

The consideration of comparative cognition is interesting because it leads naturally to a closer inspection of the Un-Cartesian position. It is certainly true that animals ‘think’ in the sense that their behaviour is, in many cases, best explained by recourse to the postulation of ‘physical symbol systems’ (Gallistel, 2009). However, this is distinct from the conception of ‘thinking’ relevant to the Un-Cartesian position. This is captured by the biconditional:

Intentionality ↔ intensionality ↔ reference ↔ grammar (Hinzen and Sheehan, 2013: 108)<sup>146</sup>

---

<sup>145</sup> As well as grammar, the vast disparity between the size of a normal human lexicon and that of even the most rigorously language-trained ape indicates that lexical capability is another area of linguistic discontinuity (Pinker and Jackendoff, 2005: 7).

<sup>146</sup> Hinzen and Sheehan also include as part of this complex biconditional, the condition of ‘concepts’. I omit this here not because I consider this mistaken or inconsistent with the account I offer, but simply because, from the present perspective, it overcomplicates matters and so obscures the central point.

The Un-Cartesian position is that grammar is intimately related to ‘thinking’ which is *intentional, intensional* and *referential*. A few clarifications are in order. Firstly, what ‘reference’ denotes over and above ‘intentionality’ eludes me, and so I take these to be equivalent concepts. Secondly, Hinzen and Sheehan stress that intentionality is a matter of states or representations being about facets of the world which exist distinctly and independently of their being so represented. In other words, the relevant concept of intentionality is bound up with the possibility of mistake or error (Hinzen and Sheehan, 2013: 463-4).

The Un-Cartesian view is that thought of this kind depends upon grammar. Grammar functions so as to give us “different kinds of thoughts to think”: namely, thoughts bearing semantic properties (Hinzen and Sheehan, 2013: 406). As a consequence, grammar cannot be ‘mere syntax’, but is of inherent semantic significance.<sup>147</sup> For instance, Hinzen and Sheehan (2013: 161) bemoan how the assumption that semantic content is independent of grammar has been so historically influential that “it is by now firmly institutionalised, with philosophers focusing solely upon ‘content’ ... while syntacticians are involved with ‘pure form’”. For the Un-Cartesian, this institutional division of labour is founded upon, and serves to perpetuate, an intellectual confusion.

It is worth underlining the striking resemblance between the ‘thought’ or ‘cognition’ targeted by the Un-Cartesian and Kant’s conception of objective cognition or experience. In my view, Kantian experience and the intentional and intensional ‘thinking’ of the Un-Cartesian overlap, in their essential elements, completely. The Un-Cartesian claim then is that thinking of this specific kind, *objective cognition*, depends upon possessing a grammatical competence.

Now, while it is certainly true that animals ‘think’ in the generic sense of manipulating symbols computationally, compellingly, they exhibit no evidence of objective cognition. The most striking feature of primate calls in their natural habitats is their functional character. In particular, they function so as to influence the behaviour of others in a limited number of (identifiable) ways e.g. alarm calls, requests for food and mating cries

---

<sup>147</sup> Hinzen and Sheehan (2013: 491) point out that Chomsky’s famous example ‘colourless green ideas sleep furiously’, which has been almost universally interpreted as evidence in favour of a grammar-meaning divide, actually supports no such divide. The sentence is grammatical, but also perfectly meaningful: we can assign it truth conditions unproblematically. By contrast, true examples of meaningless sentences, e.g. ‘furiously sleep ideas green colourless’, are also ungrammatical. The domain of grammaticality and the domain of meaning intersect completely.

(Hauser, et al, 2002: 9). For example, toque macaque monkeys make calls when encountering an abundant food source which quickly brings their group mates to the site. Vervet monkeys use three different alarm calls to alert others to the presence of eagles, leopards and snakes (Tomasello, 1997: 250-254). This functional character is reinforced by how animal calls are prompted by objects or events “in the here and now” (Hauser et al, 2007: 105). It follows from this that we need not ascribe any semantics or referential properties to these calls; they can plausibly be analysed in purely functional, that is, non-intentional, terms

This conclusion is equally viable for language-trained, signing chimpanzees. There are numerous cases of chimpanzees being trained to produce signs in a variety of contexts and these signs have often been analysed as bearing semantic properties: designating things in the world. However, similarly to primate calls in their natural environment, the signs that language trained apes produce are centred upon a narrow range of requests e.g. ‘eat’, ‘tickle’ and ‘drink’, i.e. attempts to influence the behaviour of the human experimenter. Indeed, of the productions of such chimpanzees, more than 95% of them are requests (Call and Tomasello, 1997: 323). Not all of the signs of a language-trained chimpanzee’s vocabulary are overtly functional in this way. For example Premack trained a chimpanzee, Sarah, to use an artificial visual language consisting of plastic chips of different colour shapes. In a study, Sarah would be asked “what colour?” in the presence of an apple, to which she consistently selected the symbol for ‘red’. Likewise, if asked “what object?” in the presence of an apple, she would select the symbol for ‘apple’. Crucially though, a successful answer would be rewarded with food. Terrace argues that there is no evidence that the symbols used by apes have “any function other than to request whatever food or drink its trainer could provide” (Terrace, 2005: 15). Terrace contends that if the trainer stopped offering rewards, Sarah would simply stop answering the questions.

It is plausible therefore that animal communicative acts do not bear semantic or referential properties. Rather than being ‘about’ anything in the world, they are better characterised functionally. In absence of referential utterances however, there seems to be little reason to ascribe animals intentional thought. As Terrace (2005: 18) conjectures, referential ability appears to be uniquely human.

Far from comparative cognition supplying evidence against the Un-Cartesian position then, Hinzen and Sheehan argue that it indicates exactly the kind of negative correlation - no grammar, no intentionality - which the Un-Cartesian would predict. Moreover, they contend

that this conditional relation is further supported by the absence of grammar-intentional thought dissociations or selective breakdowns in humans (2013: Chpt. 8). Firstly, it is argued that since so-called agrammatic aphasia is best analysed as a breakdown of morpho-phonology rather than a breakdown of grammar, there is no convincing evidence of selective or isolated grammatical impairment, contrary to what is widely supposed.<sup>148</sup> Secondly, Hinzen and Sheehan argue that where intentional cognition is disordered, such as it is for the schizophrenic, this plausibly correlates with a breakdown of grammar: “where language fragments, [intentional] thought does too” (Hinzen and Sheehan, 2013: 477). In sum, the basic picture is that where there is intentional cognition, there is grammar, and where there is no grammar or an impairment of grammar, intentional cognition is disrupted accordingly.

Clearly, this does not amount to a comprehensive assessment of the Un-Cartesian position. Indeed, such an assessment lies altogether beyond the reach of this thesis: it is not something I attempt to broach here. Rather, my aim is to present the Un-Cartesian view as a viable conception of grammar and, crucially, underline its relation to the central proposal of this chapter: that grammatical theory is capable of supplying the forms of judgement. At this stage it is already fairly transparent that the Un-Cartesian conception of grammar, connected as it is with objective cognition, is likely to be conducive to and motivate the proposal that the forms of judgement can be derived from the forms of grammar. Nevertheless, for a precise formulation of this relation, it is necessary to first review another source of evidence for the Un-Cartesian position.

Perhaps, the most important argument in favour of the Un-Cartesian view is that grammatical categories systematically correlate with semantic categories. For instance, in most linguistic contexts proper names must take wide scope over modal operators, strongly indicating that they are directly referential expressions.<sup>149</sup>

(1) Alfred might have had a different mother.

Nevertheless, in other linguistic contexts, proper names behave differently:

(2) The Alfred who joined the club today was a baboon.

(3) Some Alfreds are crazy; some are sane.

---

<sup>148</sup> For instance, Jackendoff (1996) cites the purported existence of double dissociations as an argument in favour of the separation of language and thought.

<sup>149</sup> Of course, this is not uncontroversial (see, e.g. Matushansky, 2008), but since the debate regarding the semantics of proper names suffers from serious confusions (Kirkby, 2012), I will avoid it altogether here.

Burge (1973) notes how in these cases the proper name is predicative rather than directly referential. However, what has not always received attention is that the difference between (1) on the one hand, and (2) and (3) on the other corresponds to a grammatical difference also:

[<sub>DP</sub> Alfred<sub>1</sub> [<sub>NP</sub> t<sub>1</sub>]] might have had a different mother

[<sub>DP</sub> The [<sub>NP</sub> Alfred]] who joined the club today was a baboon

[<sub>DP</sub> Some [<sub>NP</sub> Alfreds]] are crazy; some are sane

Indeed, it is possible that *every* grammatical distinction corresponds to a semantic one (Hinzen and Sheehan, 2013). However, observing systematic correlations between grammatical and semantic categories is one thing (Carlson, 2003; Longobardi, 2005), explaining them is another. It is the Un-Cartesian position which, perhaps uniquely, is capable of such an explanation, and this constitutes in argument in its favour. According to the Un-Cartesian, grammatical and semantic distinctions coincide because grammatical distinctions *are* semantic distinctions.

“...grammar *narrowly constrains* the ways in which words can be used to refer, making available a small number of discrete options in which this can happen, ranging from purely predicative nominals to quantificational, to referential, to deictic, and finally, to personal ones” (Hinzen and Sheehan, 2013: 211)

By means of illustration, consider the following string of examples:

(1) a. John is **man** enough to solve this problem.

b. John is more **man** than boy.

(2) John likes **man**.

(3) John is a **man**.

(4) John likes a **man**.

(5) John likes **men**.

(6) a. John is looking for **men**.

b. John is looking for three **men**.

(7) John is the **man**.

(8) **Man** arrived.

(9) John is this **man**.

In these examples, ‘man’ functions, in semantic terms, rather differently. According to Hinzen and Sheehan (2013: 218), it refers variously to an abstraction (1a & 1b), an arbitrary amount of kind (2), an arbitrary number of individual instances of a kind (6a), a set of sets with a particular cardinality (6b), a kind (5, 8), a property (3 & 7), a particular instance of a kind potentially unknown to the speaker (4) and a particular individual instance of a kind known to the speaker (9). The question is how to explain these semantic differences: the Un-Cartesian explains them grammatically.

Crucially, there is clearly something *common* to each of these cases: a lexical item or lexical root, *man*, which provides a certain descriptive content. The alternate option would be to say that distinct lexical items are being utilised in each of these examples. Yet this would entail an explosion in the size of a subject’s lexicon. Moreover, it would not account for the productivity of these examples. Any speaker of English who can comprehend one of these examples can comprehend the others also. So, there is a common lexical root underlying each of these examples, yet they also differ in their referential or semantic properties. This difference, claims the Un-Cartesian, is explained by the grammatical configurations the lexical item enters into. Specifically, the referential properties of the expression are determined by grammatical factors such as whether an edge is projected, whether the edge is light or heavy, whether the content of the interior is moved to the edge, or whether the edge is expanded by additional layers (Hinzen and Sheehan, 2013: 299)

As should be clear, the point here is not merely that certain referential properties *correlate* with grammatical configurations. The claim is rather that these configurations *explain* the relevant semantic properties. It is the function of grammar to ‘take’ semantically under-determined lexical roots such as *man*, *run*, *love*, and fix their relation to an object according to a particular referential mode. In this sense, grammar constrains any act of reference, demarcating a space of referential possibilities which cannot be circumvented. To borrow a metaphor from Hinzen (2009: 128), on this view language is not the ‘dress’ of cognition, a mere guise or clothing in which it is presented, but the *skeleton* of cognition.

Moreover, this space of referential possibilities cannot be rationalised by any non-grammatical means:

“One might formalize them as semantic or logical ones in some available notation – but this won’t change the fact that they have their basis in the structure of human

language [grammar], where they are empirically manifest.” (Hinzen and Sheehan, 2013: 408)

Why grammar imposes the constraints it does upon reference is, claim Hinzen and Sheehan, essentially a mystery. Specifically, it cannot be answered by recourse to some non-grammatical domain of ‘semantics’ or logic or metaphysics. In this sense, grammar is as deep as it is possible to go.<sup>150</sup>

Now, I have here only provided a glimpse of the domain of grammatical semantics postulated by the Un-Cartesian, a domain which encompasses clausal distinctions as well as the nominal distinctions outlined above. Still though, at this stage matters are sufficiently clear, I think, to state precisely why the Un-Cartesian conception of grammar is so conducive to the proposal that the forms of judgement be updated in light of grammatical theory. The Un-Cartesian view is not merely that grammar is, in some generic sense, connected with or required for objective cognition, as noted earlier, but more specifically that it imposes non-trivial *constraints* upon such cognition. Hinzen and Sheehan characterise these referential modes or constraints in terms of a formal ontology. They control for the type of external object posited (2013: 176), or provide “perspectives on reality” (2013: 141). Roughly, grammar configures all relations to objects in terms of a limited number of ontological categories.

However, it is possible, I think, to dispense with this ontological framework and, in a Kantian vein, understand grammatical semantics instead simply as rules, some selection of which must be realised by any judgement. That is, they can be understood as *meta-concepts*. For instance, ‘John is man enough to solve his problems’ and ‘John is a man’ have different inferential profiles, in part, because ‘man’ in the former case refers to an abstraction (a gradable property), in the latter case refers to a non-gradable property. Insofar as this is a *grammatical* phenomenon, an instance of grammatical semantics, it makes perfect sense to say that grammar provides rules for thinking about objects. Furthermore, insofar as some selection of these rules must be realised by any referential act, as the Un-Cartesian proposes, they must be known a priori: presupposed by objective experience, rather than deriving from it. As such, they are plausibly a priori rules for cognition of objects; foundational rules which

---

<sup>150</sup> This is comparable to Kant’s view that the pure forms of intuition and of the intellect are ‘mysterious’, in the sense that they cannot be rationalised (B145-6).



must be applicable to objects insofar as they are cognisable at all. That is, the rules of grammar may just be intellectual conditions for the possibility of experience

## Conclusion

---

I said in the introduction to this thesis that rather than provide a theory of judgement, an ill-conceived endeavour I suggested, I would concern myself with two distinct conceptions of judgement, seeking to (a) to locate them within the context of the relevant philosophical and scientific undertakings and, (b) show how they are mediately related via the problem of intentionality. In developing this primary thread however, I have made a number of ‘supplementary claims’, which are philosophically significant in their own right, concerning, for example, the naturalistic credentials of Hume’s theory of ideas, the connection between Hume’s science of man and cognitive science and the status of Kant’s forms of judgement, and I shall recount these here also.

Beginning with Hume’s science of man, I explained how integral to this tradition is a specific conception of judgement. On this view, judgement is understood as a species of behavioural phenomena, one which a ‘scientist of man’ needs to explain. Subjects’ judgements are the data-set or explananda in need of naturalistic explanation. For Hume, this explanation is a matter of tracing the causal genesis of the judgement back to a domain of primitive mental elements and relations. This prompted the question of whether Hume’s theory of ideas is naturalistically acceptable. I argued that, contrary to what is widely claimed, Hume’s science of man is perfectly consistent with his theory of ideas, and that the latter can be understood as the kind of ‘inference to the best explanation’ central to scientific inquiry. Furthermore, I elaborated upon the often made observation that Hume’s science of man is, in some sense, the intellectual forerunner of contemporary cognitive science. I sought to substantiate this observation in detail, contending, for example, that cognitive scientists adopt the same conception of judgement as Hume.

However, despite the explanatory successes this conception of judgement has enabled, I contended that Hume encounters a major problem regarding the intentionality of judgements. In particular, he is unable to account for a central property of his own explananda: that such judgements *represent* the world as being a certain way. Although Hume does attempt to tackle this problem when he asks why it is that we form beliefs about bodies which exist distinctly and continuously, I contended that he never fully grasps the

depth and pervasiveness of the problem, and that, besides, his proposed solution is a comprehensive failure.

By contrast, Kant was keenly occupied with this intentionality problem (or, in his terms, objective validity or objective cognition), and I begun Part II by situating this problem in the context of his broader transcendental enterprise. Subsequently, I discussed two pivotal claims made by Kant:

1. Intentional experience is essentially rule-subjecting.
2. Rules only exist in the context of a judgement (the priority principle).

As I underlined, it follows from these claims that intentional experience is necessarily judgemental and in this sense, Kant's conception of judgement is inherently bound up with the role it plays in accounting for the possibility of intentionality.

Emphasising the centrality of judgement to intentional experience prompted the question of how the forms of judgement posited by Kant ought to be understood. I argued that although Kant is right to consider these forms to be significant, he is wrong to take them to comprise a general logic. Specifically Kant is wrong to think that the forms of judgement concern mere 'well-formedness' in abstraction from any relation to an object. On the contrary, I argued that to the extent that there are no true judgements of perception, it is necessary to view the forms of judgement as inherently functioning so as to secure a relation to an object: that is, as rules for cognising objects.

The final issue addressed in this thesis concerned how the forms of judgement are to be identified. In this regard, I sought to establish the viability and attractiveness of a research program: one according to which the forms of judgement are updated in light of, or identified by recourse to, contemporary linguistic theory. I outlined the philosophical results such a program could deliver: a naturalisation of this component of Kant's philosophy, a realisation of the kind of bi-directional relationship between philosophy and the sciences of the mind which originally motivated the thesis. I sought to bolster the credibility of this program by refuting some possible objections, and pointing out that the emerging Un-Cartesian conception of grammar is highly conducive to it. Indeed, whether the proposed update is ultimately viable or not likely depends upon the success of the Un-Cartesian view of grammar.

## References

---

- Abend, G. 2012. 'What the science of morality doesn't say about morality', *Philosophy of the Social Sciences*, XX(X): 1-44.
- Allison, H. E. 1996. *Idealism and Freedom* (Cambridge: Cambridge University Press).
- Allison, H. E. 2004. *Transcendental Idealism* (London: Yale University Press).
- Allison, H. E. 2008. *Custom and Reason in Hume* (Oxford: Clarendon Press).
- Ameriks, K. 1978. 'Kant's transcendental deduction as a regressive argument', *Kant Studien*, 69(1-4): 273-287.
- Ameriks, K. 1982. 'Recent work on Kant's theoretical philosophy', *American Philosophical Quarterly*, 19(1): 1-24.
- Anderson, L. R. 2008. 'Comments on Wayne Martin, *Theories of Judgement*', *Philosophical Studies*, 137: 91-108.
- Arnauld, A. 1662. *The Art of Thinking* (Indianapolis: The Bobbs-Merrill Company).
- Baker, M. 2001. *The Atoms of Language: The Mind's Hidden Rules of Grammar* (New York: Basic Books).
- Baker, M. 2008. 'The creative aspect of language use and nonbiological nativism', in P. Carruthers, S. Lawrence & S. Stich (eds.), *The Innate Mind, Vol. 3* (Oxford: Oxford University Press), pp. 233-253.
- Bell, D. 1979. *Frege's Theory of Judgement* (New York: Oxford University Press).
- Bell, D. 2001. 'Some kantian thoughts on propositional unity', *Proceedings of the Aristotelian Society, Supplementary Volume*, 75:1-16.
- Bennett, J. 1966. *Kant's Analytic* (London: Cambridge University Press).
- Berkeley, G. 1710. *The Principles of Human Knowledge* (Fontana, 1962).

- Bermudez, J. L. 2010. *Cognitive Science* (Cambridge: Cambridge University Press).
- Berwick, R. C. & N. Chomsky 2011. 'The biolinguistic program: the current state of its development', in A. M. Di Sciullo & C. Boeckx (eds.), *The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty* (Oxford: Oxford University Press), pp. 19-41.
- Bird, G. 1962. *Kant's Theory of Knowledge* (London: Routledge & Kegan Paul).
- Biro, J. 1993. 'Hume's new science of the mind' in D. F. Norton (ed.), *The Cambridge Companion to Hume* (Cambridge: Cambridge University Press), pp. 33-63.
- Blackburn, S. Manuscript. *Review of Pinker 2002*. Retrieved from: <http://www.phil.cam.ac.uk/~swb24/reviews/Pinker.htm>
- Bloomfield, L. 1933. *Language* (New York: Henry Holt).
- Boeckx, C. 2011. 'Approaching parameters from below', in A. M. Di Sciullo & C. Boeckx (eds.), *The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty* (Oxford: Oxford University Press), pp. 205-221.
- Borg, E. 2004. *Minimal Semantics* (Oxford: Oxford University Press).
- Brandom, R. 1994. *Making it Explicit* (Cambridge, MA: Harvard University Press).
- Brandom, R. 2002. *Tales of the Mighty Dead* (Cambridge, MA: Harvard University Press).
- Brandom, R. 2009. *Reasons in Philosophy: Animating Ideas* (Harvard University Press).
- Burge, T. 1973. 'Reference and proper names', *The Journal of Philosophy*, 70(14): 425-439.
- Burgess, J. 2009. *Philosophical Logic* (New Jersey: Princeton University Press).
- Burton-Roberts, N. 2011. 'On the grounding of syntax and the role of phonology in human cognition', *Lingua*, 121(14): 2089-2102.
- Cabrera, M. 2011. 'Natural unity and human exceptionalism', *Midwest Studies in Philosophy*, XXXV: 46-68.
- Call, J. & M. Tomasello 1997. *Primate Cognition* (Oxford: Oxford University Press).
- Carlson, G. 2003. 'Interpretive asymmetries in major phrases', in A.

DiSciullo (ed.), *Asymmetry in Linguistic Theory* (Benjamins), pp. 299-311.

Cassirer, E. 1953. *Substance and Function* (New York: Dover).

Cheney, D. L. & R. M. Seyfarth 1997. 'Why animals don't have language', in G. B. Pearson (ed.), *The Tanner Lectures on Human Values* (Salt Lake City: University of Utah Press), pp. 175-209.

Chisholm, R. M. 1957. *Perceiving: A Philosophical Study* (London: Oxford University Press).

Chomsky, N. 1959. 'Review of Skinner 1957', *Language*, 35(1): 26-58.

Chomsky, N. 1966. *Cartesian Linguistics* (New York: Harper and Row).

Chomsky, N. 1980. *Rules and Representations* (New York: Columbia University Press).

Chomsky, N. 1981. *Lectures on Government and Binding* (Berlin: Mouton de Gruyter).

Chomsky, N. 1986. *Knowledge of Language: Its Nature, Origins and Use* (New York: Praeger).

Chomsky, N. 1988. *Language and Problems of Knowledge* (Cambridge, MA: MIT Press).

Chomsky, N. 2000. *New Horizons in the Study of Language and Mind* (Cambridge: Cambridge University Press).

Chomsky, N. 2005. 'Three factors in language design', *Linguistic Inquiry*, 36(1): 1-22.

Chomsky, N. 2007. 'Biolinguistic explorations: design, development, evolution', *International Journal of Philosophical Studies*, 15(1): 1-21.

Chomsky, N. 2007a. 'Approaching UG from below', in U. Sauerland, & H. M. Gartner (eds.), *Interfaces + Recursion = Language?: Chomsky's Minimalism and the View from Syntax-Semantics* (Berlin: Mouton de Gruyter), pp. 1-29.

Cinque, G. 1999. *Adverbs and Functional Heads* (Oxford: Oxford University Press).

Collins, J. 2004. 'Faculty disputes', *Mind and Language*, 19(5): 503-533.

Collins, J. 2008. *Chomsky: A Guide for the Perplexed* (Continuum International Publishing).

- Crain, S. & P. Pietroski 2005. 'Innate ideas', in J. McGilvray (ed.), *The Cambridge Companion to Chomsky* (Cambridge: Cambridge University Press), pp. 164-181.
- Cresswell, M. 2001. 'Modal logic' in L. Goble (ed.), *The Blackwell Guide to Philosophical Logic* (Blackwell Publishing).
- Cudworth, R. 1731. *A Treatise Concerning Eternal and Immutable Morality* (Cambridge: Cambridge University Press, 1996).
- Cushman, F., L. Young & J. Greene 2010. 'Our multi-system moral psychology: towards a consensus view' in J. Doris, G. Harman, S. Nichols, J. Prinz, W. Sinnott-Armstrong & S. Stich (eds.), *The Oxford Handbook of Moral Psychology* (Oxford: Oxford University Press).
- Cushman, F., L. Young & M. Hauser 2006. 'The role of conscious reasoning and intuition in moral judgement', *Psychological Science*, 17(12): 1082-1089.
- Davidson, D. 1982. 'Rational animals', *Dialectica*, 36(4): 318-327.
- Davidson, D. 1999. 'The emergence of thought', *Erkenntnis*, 51: 7-17.
- Davidson, D. 2004. 'What thought requires', in D. Davidson (ed.), *Problems of Rationality* (Oxford: Oxford University Press), pp. 136-150.
- Descartes, R. 1952. *Rules for the Direction of the Mind* (London: William Benton).
- Descartes, R. 1637. *Discourse on the Method for the Conducting One's Reason Well and for Seeking the Truth in the Sciences* (London: William Benton, 1952).
- Descartes, R. 1644. *The Principles of Philosophy* (Kessinger Publishing).
- Dickersen, A. B. 2004. *Kant on Representation and Objectivity* (Cambridge: Cambridge University Press).
- Dreyfus, H. L. 2007. 'The return of the myth of the mental', *Inquiry*, 50(4): 352-365.
- Dupoux E. & P. Jacob 2007. 'Universal moral grammar: a critical appraisal', *Trends in Cognitive Science*, 11(9): 373-379.

- Dwyer, S. 2007. 'How good is the linguistic analogy?', in P. Carruthers, S. Laurence & S. Stich (eds.), *The Innate Mind: Volume 2: Culture and Cognition* (Oxford: Oxford University Press), pp. 238-256.
- Dwyer, S. 2009. 'Moral dumfounding and the linguistic analogy: methodological implications for the study of moral judgement', *Mind and Language*, 24(3): 274-296.
- Dwyer, S., B. Huebner & M. Hauser 2010. 'The linguistic analogy: motivations, results and speculations', *Topics in Cognitive Science*, 2: 486-510.
- Everett, D. 2005. 'Cultural constraints upon grammar and cognition in Piraha', *Current Anthropology*, 4(46): 621-646.
- Fitch, T. & M. Hauser 2004. 'Computational constraints on syntactic processing in a nonhuman primate', *Science*, 303: 377-380.
- Flage, D. E. 1990. *David Hume's Theory of Mind* (London: Routledge).
- Flew, A. 1986. *David Hume* (Oxford: Basil Blackwell Inc.).
- Fodor, J. 1975. *The Language of Thought* (New York: Thomas Y. Crowell Company).
- Fodor, J. 1983. *The Modularity of the Mind* (Cambridge, MA: MIT Press).
- Fodor, J. 2001. 'Language, thought and compositionality', *Mind and Language*, 16(1): 1-15.
- Fodor, J. 2003. *Hume Variations* (Oxford: Oxford University Press).
- Fodor, J. 2007. 'The revenge of the given', in B. P. McLaughlin & J. D. Cohen (eds.), *Contemporary Debates in Philosophy of Mind* (Oxford: Blackwell), pp. 105-116.
- Fodor, J. & Z. Pylyshyn 1988. 'Connectionism and cognitive architecture: a critical analysis', *Cognition*, 28: 3-71.
- Foot, P. 1978. 'The problem of abortion and the doctrine of double effect', in *Virtues and Vices and Other Essays in Moral Philosophy* (Oxford: Basil Blackwell).
- Frege, G. 1897. 'Logik', in H. Hermes, F. Kambartel & F. Kaulbach (eds.), *Posthumous Writings* (Oxford: Blackwell, 1979), pp. 126-152.
- Frege, G. 1956. 'The thought: a logical inquiry', *Mind*, New Series, 65(259): 289-311.



- Gallistel, C. R. 2009. 'The foundational abstractions', in M. Piattelli-Palmarini, J. Uriagereka & P. Salaburu (eds.), *Of Minds and Language* (Oxford: Oxford University Press), pp. 58-73.
- Garrett, D. 1997. *Cognition and Commitment in Hume's Philosophy* (Oxford: Oxford University Press).
- Garrett, D. 2006. 'Hume's naturalistic theory of representation', *Synthese*, 152: 301-319.
- Ginsborg, H. 2008. 'Was Kant a nonconceptualist?', *Philosophical Studies*, 137: 65-77.
- Glock, H., 2010. 'Can animals judge?', *Dialectica*, 64(1): 11-33.
- Gomila, A. 2012. *Verbal Minds* (London: Elsevier).
- Gray, K., A. Watz & L. Young 2012. 'Mind perception is the essence of morality', *Psychological Inquiry*, 23: 101-124.
- Green, T. H. 1874. 'Introduction to Hume's Treatise of Human Nature', in R. L. Nettleship (ed.), *Works of Thomas Hill Green* (Cambridge: Cambridge University Press, 2011).
- Greenberg, J. H. 1963. *Universals of Language* (London: MIT Press).
- Greene, J., L. Nystrom, A. Engell, J. Darley & J. Cohen 2004. 'The neural bases of cognitive conflict and control in moral judgement', *Neuron*, 44: 389-400.
- Grene, M. 1994. 'The objects of Hume's Treatise', *Hume Studies*, XX(2): 163-177.
- Guyer, P. 1987. *Kant and the Claims of Knowledge* (Cambridge: Cambridge University Press).
- Haidt, J. 2001. 'The emotional dog and its rational tail: a social intuitionist approach to moral judgment', *Psychological Review*, 108(4): 814-834.
- Haidt, J. & C. Joseph 2007. 'The moral mind: how 5 sets of innate moral intuitions guide the development of many culture-specific virtues, and perhaps even modules', in P. Carruthers, S. Laurence & S. Stich (eds.), *The Innate Mind, Vol. 3*. (New York: Oxford University Press), pp. 367-391.
- Haidt, J., S. Koller & M. Dias 1993. 'Affect, culture, and morality, or is it wrong to eat your dog?', *Journal of Personality and Social Psychology*, 65: 613-628.

- Hanna, R. 2001. *Kant and the Foundations of Analytic Philosophy* (Oxford: Oxford University Press).
- Hanna, R. 2005. 'Kant and nonconceptual content', *European Journal of Philosophy*, 13: 247-290.
- Hanna, R. 2011. 'Beyond the myth of the myth: a kantian theory of non-conceptual content', *International Journal of Philosophical Studies*, 19(3): 323-398.
- Harman, G. 2008. 'Using a linguistic analogy to study morality', in W. Sinnott-Armstrong (ed.), *Moral Psychology, Vol. I*, (Cambridge, MA: MIT press), pp. 345-351.
- Haspelmath, M. 2008. 'Parametric versus functional explanations of syntactic universals', in T. Biberauer (ed.), *The Limits of Syntactic Variation* (Amsterdam: Benjamins), pp. 75-107.
- Hatfield, G. 1990. *The Natural and the Normative* (Cambridge, Massachusetts: MIT Press).
- Hatfield, G. 1997. 'The workings of the intellect: mind and psychology', in P. A. Easton (ed.), *Logic and the Workings of the Mind: The Logic of Ideas and Faculty Psychology in Early Modern Philosophy* (California: Ridgeview Publishing Company), pp. 21-45.
- Hauser, M. 2006. *Moral Minds* (New York: HarperCollins).
- Hauser, M., D. Barner & T. O'Donnell 2007. 'Evolutionary linguistics: a new look at an old landscape', *Language Learning and Development*, 3(2): 101-132.
- Hauser, M., N. Chomsky & T. Fitch 2002. 'The faculty of language: what is it, who has it and how did it evolve?', *Science*, 298: 1569-1579.
- Heis, J. Forthcoming. 'The priority principle from Kant to Frege', *Nous*.
- Hinzen, W. 2006. *Mind Design and Minimal Syntax* (Oxford: Oxford University Press).
- Hinzen, W. 2007. *An Essay on Names and Truth* (Oxford: Oxford University Press).
- Hinzen, W. 2009. 'Hierarchy, merge and truth', in M. Piattelli-Palmarini, J. Uriagereka & P. Salaburu (eds.), *Of Minds and Language* (Oxford: Oxford University Press), pp. 123-141.
- Hinzen, W. 2012. 'The philosophical significance of universal grammar', *Language Sciences*, 34: 635-649.

- Hinzen, W. 2012a. 'Minimalism', in R. Kempson, T. Fernando & N. Asher (eds.), *Philosophy of Linguistics* (Oxford: North Holland), pp. 93-143.
- Hinzen, W. & M. Sheehan 2013. *The Philosophy of Universal Grammar* (Oxford: Oxford University Press).
- Howell, R. 1992. *Kant's Transcendental Deduction* (Dordrecht: Kluwer Academic Publishers).
- Huebner, B. 2011. 'Critiquing empirical moral psychology', *Philosophy of the Social Sciences*, 41(1): 50-83.
- Hume, D. 1739. *A Treatise of Human Nature* (New York: Barnes and Noble, 2005).
- Hume, D. 1748. *An Enquiry Concerning Human Understanding* (Oxford: Oxford University Press, 1999).
- Jackendoff, R. 1996. 'How language helps us think', *Pragmatics and Cognition*, 4(1): 1-34.
- Jackendoff, R. 2002. *Foundations of language* (Oxford: Oxford University Press).
- Joos, M. 1966. *Readings in Linguistics*, 4<sup>th</sup> ed. (Chicago: University of Chicago Press).
- Joyce, R. 2007. *The Evolution of Morality* (Cambridge, MA: MIT Press).
- Kant, I. 1762. 'The false subtlety of the four syllogistic figures', in D. Walford (ed.), *Theoretical Philosophy 1755-1700* (Cambridge: Cambridge University Press, 1922), pp. 85-105.
- Kant, I. 1781. *Critique of Pure Reason* (Cambridge: Hackett Publishing Company, 1996).
- Kant, I. 1783. *Prolegomena to Any Future Metaphysics That Will Be Able to Come Forward as Science* (London: George Bell and Sons, 1909).
- Kant, I. 1786. *Metaphysical Foundations of Natural Science* (London: George Bell and Sons, 1909).
- Kant, I. 1800. *Logic* (Indianapolis: Bibbs-Merrill Company Inc., 1974).
- Kant, I. 1999. *Correspondence* (Cambridge: Cambridge University Press).

Katz, J. & D. Pesetsky Manuscript. *The Identity Thesis for Language and Music*. Retrieved from:

<http://ling.auf.net/lingBuzz/000959>

King, J. 2009. 'Questions of unity', *Proceedings of the Aristotelian Society*, CIX(3): 258-277.

Kirkby, D. 2012. 'Proper names: denying semantic uniformity', in P. Stalmaszczyk (ed.), *Philosophical and Formal Approaches to Linguistic Analysis* (Frankfurt: Ontos), pp. 263-273.

Kirkby, D. 2013. 'Why there might be a moral faculty', *Philosophical Psychology*. 1-8.

Kirkby, D. & U. Reichard Manuscript. *Developing the Linguistic Analogy*.

Kirkby, D., W. Hinzen & J. Mikhail 2013. 'Your theory of the evolution of morality depends upon your theory of morality', *Behavioral and Brain Sciences*, 36: 94-95.

Kitcher, P. 1990. *Kant's Transcendental Psychology* (Oxford: Oxford University Press).

Knobe, J. 2007. 'Experimental philosophy and philosophical significance', *Philosophical Explorations*, 10(2): 119-121.

Kohlberg, L. 1981. *Essays on Moral Development, Volume I: The Philosophy of Moral Development* (San Francisco, CA: Harper & Row).

Krebs, D. & K. Denton 2005. 'Toward a more pragmatic approach to morality: a critical evaluation of Kohlberg's model', *Psychological Review*, 112(3): 629-649.

Kriegel, U. 2008. 'Real narrow content', *Mind and Language*, 23(3): 304-328.

Kukla, T. A. 2011. *Kant's Theory of Cognition: An Interpretation of the Argument of the Transcendental Deduction*, Unpublished PhD, University of Illinois.

Landy, D. 2008. 'Sellars on Kant and Hume on representing complexes', *European Journal of Philosophy*, 17(2): 224-246.

Landy, D. 2009. 'Inferentialism and the transcendental deduction', *Kantian Review*, 14(1): 1-30.

Landy, D. Manuscript. *From Sensations to Intuitions: Kant on Non-Conceptual Content*.

- Leibniz, G. W. 1696. 'Letter to Gabriel Wagner on the value of logic', in L. Loemker (ed.), *Philosophical Papers and Letters* (London: Kluwer, 1969), pp. 462-472.
- Lettvin, J. Y., H. R. Maturana, W. S. McCulloch & W. H Pitts 1959. 'What the frog's eye tells the frog's brain', *Proceedings of the Institute of Radio Engineers*, 47.
- Locke, J. 1690. *An Essay Concerning Human Understanding* (Oxford: Clarendon Press, 1924).
- Longobardi, G. 2005. 'Toward a unified grammar of reference', *Zeitschrift für Sprachwissenschaft*, 24(1): 5-44.
- Longuenesse, B. 1998. *Kant and the Capacity to Judge* (New Jersey: Princeton University Press).
- Macfarlane, J. 2000. *What does it Mean to Say that Logic is Formal?*, Unpublished PhD, University of Pittsburg.
- Macfarlane, J. 2002. 'Frege, Kant and logic in logicism', *The Philosophical Review*, 111(1): 25-65.
- Mallon, R. & S. Nichols 2010. 'Rules', in J. Doris & The Moral Psychology Research Group (eds.), *The Moral Psychology Handbook* (Oxford: Oxford University Press), pp. 297-320.
- Manzini, M. R. & K. Wexler 1987. 'Parameters, binding theory and learnability', *Linguistic Inquiry*, 18(3): 413-444.
- Marr, D. 1982. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information* (New York: Freeman).
- Martin, W. 2006. *Theories of Judgement* (Cambridge: Cambridge University Press).
- Matushansky, O. 2008. 'On the linguistic complexity of proper names', *Linguistics and Philosophy*, 21: 573-627.
- McDowell, J. 1994. *Mind and World* (Cambridge, MA: Harvard University Press).
- McDowell, J. 2009. *Having the World in View* (Cambridge, MA: Harvard University Press).
- McLear, C. 2011. 'Kant on animal consciousness', *Philosophers Imprint*, 11(15): 1-16.

Melnick, A. 1993. 'Review of Kant's *Transcendental Psychology*', *Philosophy of Science*, 60(3): 513-515.

Mikhail, J. 2000. *Rawls's Linguistic Analogy: A Study of the 'Generative Grammar' Model of Moral Theory Described by John Rawls in 'A Theory of Justice'*, Unpublished PhD, Cornell University, NY.

Mikhail, J. 2007. 'Universal moral grammar: theory, evidence and the future', *Trends in Cognitive Science*, 11(4): 143-152.

Mikhail, J. 2009. 'Moral grammar and intuitive jurisprudence: a formal model of unconscious moral and legal knowledge', in D. Medin, L. Skitka, C. W. Bauman & D. Bartels, (eds.), *The Psychology of Learning and Motivation: Moral Cognition and Decision Making* (Academic Press), pp. 28-100.

Mikhail, J. 2011. *Elements of Moral Cognition: Rawls' Linguistic Analogy and the Cognitive Science of Moral and Legal Judgement* (New York: Cambridge University Press).

Mounce, H. O. 1999. *Hume's Naturalism* (London: Routledge).

Mukherji, N. 2010. *The Primacy of Grammar* (Cambridge, MA: MIT Press).

Newmeyer, F. J. 1986. 'Has there been a 'chomskyan revolution' in linguistics?', *Language*, 62(1): 1-18.

Newmeyer, F. J. 2005. *Possible and Probable Languages: A Generative Perspective on Linguistic Typology* (Oxford: Oxford University Press).

Nichols, S. 2005. 'Innateness and moral psychology', in P. Carruthers, S. Laurence & S. Stich (eds.), *The Innate Mind, Vol. 3* (New York: Oxford University Press).

Norton, D. F. 1982. *David Hume* (New Jersey: Princeton University Press).

Norton, D. F. 1993. 'Introduction to Hume's thought', in D. F. Norton (ed.), *The Cambridge Companion to Hume* (Cambridge: Cambridge University Press), pp. 1-32.

Owen, D. 2003. 'Locke and Hume on belief, judgement and assent', *Topoi*, 22: 15-28.

Pears, D. 1990. *Hume's System* (Oxford: Oxford University Press).

- Pendlebury, M. 1995. 'Making sense of Kant's schematism', *Philosophy and Phenomenological Research*, 55(4): 777-797.
- Penn, D. C., K. J. Holyoak & D. J. Povinelli 2008. 'Darwin's mistake: explaining the discontinuity between human and nonhuman minds', *Behavioral and Brain Sciences*, 31: 109-178.
- Pereboom, D. 1988. 'Kant on intentionality', *Synthese*, 77(3): 321-352.
- Pietroski, P. 2005. 'Meaning before truth', in G. Peter & G. Preyer (eds.), *Contextualism in Philosophy* (Oxford: Oxford University Press), pp. 255-302.
- Pinker, S. 1994. *The Language Instinct* (London: Penguin Group).
- Pinker, S. 2002. *The Blank Slate: The Modern Denial of Human Nature* (London: Penguin Group).
- Pinker, S. & R. Jackendoff 2005. 'The faculty of language: what's special about it?', *Cognition*, 95: 201-236.
- Pippin, R. B. 1982. *Kant's Theory of Form* (London: Yale University Press).
- Pizarro, D. A. & P. Bloom 2003. 'The intelligence of moral intuitions: a comment on Haidt 2001', *Psychological Review*, 110(1): 193-196.
- Price, H. 2004. 'Naturalism without representation', in D. Macarthur & M. De Caro (eds.), *Naturalism in Question* (Cambridge, MA: Harvard University Press), pp. 71-88.
- Prinz, J. 2008. 'Resisting the linguistic analogy: a commentary on Hauser, Young and Cushman', in W. Sinnott-Armstrong (ed.), *Moral Psychology, Vol. 2*, (Cambridge, MA: MIT press), pp. 157-171.
- Pylyshyn, Z. 1980. 'Computation and cognition: issues in the foundations of cognitive science', *The Behavioral and Brain Sciences*, 3: 111-169.
- Quine, W. 1951. 'Two dogmas of empiricism', *The Philosophical Review*, 60(1): 20-43.
- Quine, W. 1970. *Philosophy of Logic* (New Jersey: Prentice-Hall Inc.).
- Reid, T. 1764. *An Inquiry into the Human Mind on the Principles of Common Sense* (Edinburgh: Edinburgh University Press, 1997).

- Rizzi, L. 1982. *Issues in Italian Syntax* (Dordrecht: Foris).
- Rodl, S. 2009. 'Judgment as synthesis', *European Journal of Philosophy*, 17(3): 435-446.
- de Sa Pereira, R. H. 2013. 'What is nonconceptualism in Kant's philosophy?', *Philosophical Studies*, 164(1): 233-254.
- Searle, J. R. 1983. *Intentionality* (Cambridge: Cambridge University Press).
- Sellars, W. 1953. 'Meaning and inference', *Mind*, New Series, 62(247): 313-338.
- Sellars, W. 1967. 'Language, rules and behaviour', in S. Hook (ed.), *John Dewey: Philosopher of Science and Freedom* (New York: Dial Press), pp. 289-315.
- Sellars, W. 1978. 'The role of the imagination in Kant's theory of experience', in H. W. Johnstone (ed.), *Categories: A Colloquium* (Pennsylvanian State University).
- Sheehan, M. & W. Hinzen 2012. 'Moving towards the edge', *Linguistic Analysis*, 37: 405-459.
- Shenhav, A. & J. Greene 2010. 'Moral judgments recruit domain-general valuation mechanisms to integrate representations of magnitude and probability', *Neuron*, 67: 667-677.
- Sinnott-Armstrong, W., L. Young & F. Cushman 2010. 'Moral intuitions as heuristics', in J. Doris, G. Harman, S. Nichols, J. Prinz, W. Sinnott-Armstrong & S. Stich (eds.), *The Oxford Handbook of Moral Psychology* (Oxford: Oxford University Press). Retrieved from: [http://www.mit.edu/~lyoung/Site/Publications\\_files/Sinnott-Armstrong,%20Young,%20Cushman.doc](http://www.mit.edu/~lyoung/Site/Publications_files/Sinnott-Armstrong,%20Young,%20Cushman.doc)
- Skinner, B. F. 1957. *Verbal Behaviour* (New York: Appleton-Century-Crofts Inc.).
- Sluga, H. 2008. 'Wayne Martin on judgment', *Philosophical Studies*, 137: 109-119.
- Smith, N. 1905. 'The Naturalism of Hume (I.)', *Mind*, 14(2):149-173.
- Smith, N. 1905a. 'The Naturalism of Hume (II.)', *Mind*, 14(3): 335-347.
- Stapleford, S. 2008. *Kant's Transcendental Arguments: Disciplining Pure Reason* (Continuum).



- Sterelney, K. 2010. 'Moral nativism: a critical response', *Mind and Language*, 25(3): 279-297.
- Stevenson, L. 1979. 'Recent work on the Critique of Pure Reason', *The Philosophical Quarterly*, 29(117): 345-354.
- Strawson, P. F. 1952. *Introduction to Logical Theory* (New York: John Wiley and Sons).
- Strawson, P. F. 1966. *The Bounds of Sense* (London: Methuen & Co. Ltd.).
- Strickland, B., M. Fisher & J. Knobe 2012. 'Moral structure falls out of general event structure', *Psychological Inquiry*, 23(2): 198-205.
- Stroud, B. 1977. *Hume* (Oxon: Routledge & Kegan Paul).
- Stroud, B. 2006. 'The constraints of Hume's naturalism', *Synthese*, 152(3): 339-351.
- Terrace, H. 2005. 'Metacognition and the evolution of language', in H. Terrace & J. Metcalfe (eds.), *The missing link in cognition* (Oxford: Oxford University Press), pp. 84-115.
- Travis, C. Manuscript. *Unlocking the Outer World*. Retrieved from:  
<http://mlag.up.pt/wp-content/uploads/2011/09/Unlock121.pdf>
- Van Roojen, M. 2006. 'Knowing enough to disagree', in R. Schaffer-Landau (ed.), *Oxford Studies in Metaethics: Volume 1* (Oxford: Oxford University Press), pp. 161-193.
- Warnock, G. J. 1948. 'Concepts and schematism', *Analysis*, 9: 77-82.
- Wedgwood, R. 2009. 'The normativity of the intentional', in A. Beckermann & B. P. McLaughlin (eds.), *The Oxford Handbook of the Philosophy of Mind* (Oxford: Oxford University Press), pp. 421-436.
- Whitely, C. H. 1951. 'The idea of logical form', *Mind*, New Series, 60(240): 539-541.
- Williams, M. 2004. 'The unity of Hume's philosophical project', *Hume Studies*, 30(2): 265-296.
- Winkler, K. P. 2010. 'Kant, the empiricists, and the enterprise of the deduction', in P. Guyer (ed.), *The Cambridge Companion to Kant's Critique of Pure Reason* (Cambridge: Cambridge University Press), pp. 41-73.

Wolff, R. P. 1963. *Kant's Theory of Mental Activity* (Cambridge, MA: Harvard University Press).

Xu, F. 2005. 'Categories, kinds and object individuation in infancy', in L. Gershkoff-Stowe & D. Rakison (eds.), *Building Object Categories in Developmental Time* (Hillsdale, NJ: Erlbaum), pp. 63-89.

Zangwill, N. 1998. 'Direction of fit and normative functionalism', *Philosophical Studies*, 91: 173-203.

Zangwill, N. 2010. 'Normativity and the metaphysics of mind', *Australasian Journal of Philosophy*, 88(1): 21-39.